DOCUMENT RESUME

ED 120 132 95 SP 009 904

AUTHOR Borich, Gary D.; And Others
TITLE An Evaluation of the Personalized Model of Teacher Training. Final Report.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Aug 74
CONTRACT NE-C-00-3-0066
NOTE 221p.

EDRS PRICE MF-$0.83 HC-$11.37 Plus Postage
DESCRIPTORS *Attitudes; Attitude Tests; *Effective Teaching; Higher Education; *Personality Assessment; Personality Tests; Preservice Education; *Student Teachers; Student Teaching; *Teacher Education; Video Tape Recordings
IDENTIFIERS *Personalized Teacher Education Program

ABSTRACT

This study examined the effects of the traditional and personalized approaches to teacher training as they affect teaching behavior and pupil evaluations of teaching. It also identified interactions between training approaches and the personality and attitudes of the prospective teacher. Seventy-seven teacher trainees at the University of Texas were selected to participate in the study. Thirty-nine of these students voluntarily participated in a personalized teacher education (PTE) program, while the remaining 38 students voluntarily participated in a traditional program. Both programs included university work, classroom observation, and practice teaching. The PTE program also consisted of (1) repeated counseling sessions, (2) the differential assignment of instructional tasks and activities based on the specific attitudinal and personality characteristics of the prospective teacher, (3) self-observation of teaching behavior through videotaping, and (4) affective feedback and assignment of tasks and activities related to the self-observation. Personality and attitude scales were administered to students in both programs on entry into training. Teaching effectiveness variables were measured at the end of the practice teaching semester. The program differences for observed teaching behavior and for the student teachers' perception of their training program indicated that the PTE program was producing behavior change in the desired direction. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED120132

Final Report

AN EVALUATION OF THE PERSONALIZED MODEL OF TEACHER TRAINING

Gary D. Borich
Robert C. Godbout
Robert F. Peck
Marilynn M. Kash
Lee H. Poynor

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Research and Development Center for Teacher Education

The University of Texas at Austin

August 1974

SP009 904

Final Report

An Evaluation of the Personalized Model of Teacher Training

Gary D. Borich, Project Director
Robert C. Godbout
Robert F. Peck
Marilynn M. Kash
Lee H. Poynor

This report was funded by the National Institute of Education, Contract No. NE-C-00-3-0066. The opinions expressed herein do not necessarily reflect the opinions or policy of the National Institute of Education.

Table of Contents

Section I: Introduction

Chapter 1: The Role of Personalized Teacher Education 1

Section II: Research Report

Chapter 2: Overview of the Evaluation Design 25

Chapter 3: Instrumentation 36

Chapter 4: Data Collection Procedures 76

Chapter 5: Data Analysis Strategies 89

Chapter 6: Main Effect Comparisons Between the Traditional and Personalized Models of Teacher Training 110

Chapter 7: Interactions Between Personality and Attitude Traits and the Training Programs 124

Section III: Conclusion

Chapter 8: Evaluating Personalized Teacher Training: Some Concluding Remarks 157

Appendix A 187

Appendix B 195

SECTION I.

INTRODUCTION

Chapter 1

THE ROLE OF PERSONALIZED TEACHER EDUCATION

The crisis in public schools and on college campuses came as a shocking betrayal to the American taxpayers in the late 1960's. The nostalgic memories of their own school days had been nurtured and romanticized in the decades of the 50's and 60's by popular images of teachers like "Our Miss Brooks," "Mr. Peepers," "Mr. Novak," and "The Halls of Ivy" college professors. As the images of lovable (but competent) teachers and fumbling (but knowledgeable) professors faded from public view, the nightly newscasts and daily headlines made it all too clear that the classroom and the campus had changed. In popular and professional literature the conditions and problems of education were criticized and condemned, or excused and condoned. But the problems and conditions described in such books as Crisis in the Classroom (Silberman. 1970), How to Survive in Your Native Land (Herndon, 1971), How Children Fail (Holt, 1964), Death at an Early Age (Kozol, 1967), and Growing Up Absurd (Goodman, 1960) were not just contained within the covers of books and periodicals. They were real and visible problems evidenced by barricaded buildings, striking teachers, a growing rate of school dropouts and newly organized militant minority-group parents demanding equality of education for their children.

Crisis: From Quantity to Quality

For professional educators and administrators, the crisis of the late 1960's was simply the latest and most spectacular of a series of continuing

crises that had occurred in public and private education from kindergarten through college since the Second World War. The college campuses that were exploding with violence in the 1960's had exploded with veterans in the 1940's and public school systems now crippled by striking teachers and boycotting students had bulged beyond capacity with the population growth in the aftermath of the war. The great numbers of students flowing into the educational system of this nation caused a crisis in size and space and a shortage of teachers: a crisis of quantity. The solution to that crisis was provided through the resources of a growing technological and affluent mass society. Fed by a growing suburbia, school systems multiplied and built innovative and functional educational plants. Universities and colleges of all sizes initiated multimillion dollar building programs. The institution of education reflected the massive growth of production and the growth of systemization incited by the sheer weight of numbers as did all our other national institutions of industry, government and finance.

Education in America changed remarkably in size but very little in pedagogical form or curriculum content. The role of teacher training institutions was to supply as many teachers as fast as possible through preservice and inservice programs. Certification programs based on traditional course requirements supplied the formal "quality" control along with slightly better salaries and the open job market which attracted more able and ambitious students. The net result was a massive infusion of traditionally trained teachers who met the crisis in quantity by their numbers but whose training served the existing educational system and reinforced the status quo. The effect was such that cynical critics of American education in the 1950's were saying that education was the one thing Americans seemed willing to pay for and not get.

Although professional educators were already aware of the need for changes in both curriculum and teaching methodologies, the dramatic launching of Russia's Sputnik brought the questions of curriculum adequacy and student competence into public focus. By the decade of the 60's, the curriculum reform movement was seriously and productively underway. The educational literature of this period reflects the sudden popular concern with the quality of American education. The public was informed as to why pupils could not read, who was culturally disadvantaged, what should be taught to our students, and how students should be taught. Industry and advanced technology joined with education to produce new educational "hardware," new teaching aids and computerized instructional and administrative systems. Private organizations and agencies of state and federal government solicited, sponsored and financed projects for educational research and program development in unprecedented numbers at previously unheard of costs. Quality in education would be achieved, or so it seemed, through new curricula, new methodology, new technological hardware and special remedial reading and language programs. This solution to the problem of quality depended upon three critical conditions: (1) that the reformed curricula and new methodologies would in reality improve the quality of education; (2) that new curricula and methodology could be implemented into the existing educational system; and (3) that the public would underwrite the cost of quality in education as they had the cost of quantity.

Quality Through Curriculum Reform

Curriculum reformers seriously considered all of these conditions and, with the collaboration of members of the various disciplines, had engaged in a revolutionary approach to curriculum development. According to Bruner

(1970), the resulting curricular products were "curricula that represent an extraordinary achievement in academic quality and in the respect they show for the nature of human thought processes." But the questions of implementation and the survival of a quality product when exposed to the classroom brought Bruner these observations:

> There was stress and strain when working scientists came face to face with the realities of the working teacher or the working school budget. And there were moments of despair when some of my less patient scientific colleagues talked about making their particular curriculum "teacher-proof." It was a little like making love people-proof. But even the complaints about the teacher as spoiler grew out of respect for the basic task of equipping the student with the competency inherent in the subject matter. Nothing must interfere, not even the teacher. (p. 66).

Quality Through Pedagogical Innovation

Curriculum reformers recognized the role of the teacher as implementor and were concerned with the possible weakness in the link teachers provide between new curriculum and pupils. The same concern was motivating the explosion of knowledge in the field of teaching methodology and skills where new concepts of the teaching-learning process were appearing in innovative formats. "Team teaching," "needs grouping," "self-pacing," "peer-group tutoring," "individual guidance," "instructional cycling," "contracting," and "stating behavioral objectives" became the ammunition of pedagogy with which the fortress of the old "teacher-centered classroom" would be conquered and restructured into the proper "pupil-centered" form. Alternatives to the "self-contained" classroom such as the open classroom, ungraded units, and "unwalled" schools appeared in experimental innovations throughout the country.

Unfortunately, new pedagogical concepts and methodologies are no more teacher-proof than new curricula. Many critics and professionals who have

concerned themselves with the implementation of these new concepts and methodologies in the nation's classrooms have concluded that so far the only successful achievement has been the implementation of the language of reform. Goodlad (1969) wrote:

> We were unable to discern much attention to pupil needs, attainments, or problems as a basis for individual opportunities to learn, ... Teaching was predominantly telling and questioning by the teacher, with children responding one by one or occasionally in chorus. In all of this, the textbook was the most highly visible instrument of learning and teaching . . . . Rarely did we find small groups intensely in pursuit of knowledge; rarely did we find individual pupils at work in self-sustaining inquiry . . . . we are forced to conclude that much of the so-called educational reform movement has been blunted on the classroom door. (p. 159).

Two other critics (Allen and Mackin, 1970) have concluded that we have only a facade of change in our current educational picture. This appearance of change achieved through modernistic buildings, dramatic curriculum packages and well-publicized descriptions of teaching innovations has only served to deceive both the public and professional educators. While these critics acknowledge that real change has occurred sparingly in a few sites throughout the entire country, they have also observed that there is still a sharp distinction between what is considered ideal in the classroom and what is real.

Quality: A Relevant, Personalized Education

Among pupils, parents, taxpayers and critics, the problems of our current educational crisis in quality can be summarized in three words. Education is irrelevant, depersonalized and too expensive. These criticisms have been taken seriously and responsibly by educators and professionals in other disciplines who have taken an active part in searching for viable solutions to all three problems. Alvin Toffler (1971) writes:

> One basic complaint of the student is that he is not treated as an individual, that he is served up an undifferentiated gruel, rather than a personalized product. Like the Mustang buyer, the student wants to design his own. The difference is that while industry is highly responsive to consumer demand, education typically has been indifferent to student wants. (p. 272).

The interfacing of relevance and personalization has been understood and addressed in the work of curriculum reformers. For example, Bruner (1971) has defined "relevance" in terms that clearly express an understanding of the student:

> The word has two senses. The first is that what is taught should have some bearing on the grievous problems facing the world, the solutions of which may affect our survival as a species. This is social relevance. Then there is personal relevance. What is taught should be self-rewarding by some existential criterion of being "real," or "exciting," or "meaningful." The two kinds of relevance are not necessarily the same, alas . . . . Relevance, in either of its senses, depends upon what you know that permits you to move toward goals you care about. It is this kind of "means-ends" knowledge that brings into a single focus the two kinds of relevance, personal and social. It is then that we bring knowledge and conviction together, and it is this requirement that faces us in the revolution in education through which we are going. (p. 114).

The call for relevance and personalization also struck responsive chords in the people responsible for development of instructional curricula. The appearance of the "Affective Domain" as a co-star with the "Cognitive" and "Psychomotor" domains in current instructional curricula and the inclusion of interpersonal relationship skills as well as communication skills among the required tools for teaching attest to the responsive efforts of these professionals. Inherent in all the innovative forms of teaching methodology from "self-pacing" to "peer group tutoring" is the understanding of the pupil as a feeling, contributing participator in the process.

Training Teachers as Change Agents

The hope that either curriculum reform or new teaching formats could achieve real change in our system through an educational "domino theory"

does indeed appear to be blunted at the same visible point: the classroom teacher. This "visible point" is, as one might suspect, analogous to the illustrative "tip of the iceberg." The classroom teacher rises through our total educational system, and stands as a representative of that system. It would be strange, indeed, if the public school teacher emerged from the professionally focused strata of the educational system--the teacher training institution--as a teacher who is open to innovation, adaptive to change, aware of individual needs, and steeped in the concept of the student-centered educational program. The prospective teacher may be taught about these attitudes and concepts, but in only a very few teacher training institutions are prospective teachers actually taught by these concepts and with these attitudes. The ten teacher training program models which emerged from the sponsoring efforts of the U. S. Office of Education in the last decade, while expressing unique emphases in certain aspects of their programs, all showed efforts to force teacher training programs out of the dichotomy of course content and practical experience, so that the medium, i.e., the teacher training program, becomes the message, i.e., how to teach.

The development of such model programs could lead to the erroneous conclusion that there is a known body of facts as to what constitutes effective teaching behaviors, what skills are essential to teaching and what attitudes are necessary to promote student learning. The variety, the ingenuity and the creativity of the model programs are appealing, and it would be possible to make a selection of one or another on the basis of intuition and appeal, but the truth is that these programs have no better bases than any other existing teacher education program founded on the practical experience of teachers, general psychological principles, studies in philosophy, the social sciences and respectable intentions. Rosenshine

and Furst (1971), in reviewing these models of teacher training programs, drew sharp attention to this problem:

> However, as of this writing no one has shown that the behaviors identified in the models have any proven relevance for the real world. To be real, teacher behaviors need to be researched so that they are known to have some relationship to student outcome measures. Until this research is done, we can have little confidence that the models are providing any more hope that either teacher training or student education will be greatly improved in the foreseeable future. (p. 66).

Teacher Behavior Research

It is not that research on the effects of teaching behaviors has not been attempted. Even a cursory review of the literature on this subject shows the impressive extent and volume of the research over the last three decades. But the net result of the research on teacher behavior, both past and present, has been to confirm that at this time there are no clear conclusions (Gage, 1963; Ornstein, 1971). More than one critic of research on teaching and teaching behavior has been quick to point out that there are also no generally agreed upon conceptualizations of teaching, nor have teaching behaviors or teacher characteristics been mutually identified and defined to permit any kind of generalizability or clarity in interpretation from one study to another. Categories of "good" or "effective" teachers remain descriptive and any attempt to make the "effective" behaviors of one teaching situation prescriptive for another teaching situation quite often runs headlong into reverse findings. Are there any teacher characteristics or personality traits that generalize as "good" or "effective" from one teaching situation to another? Attempts to find such generalizable characteristics have produced an unedifying descriptive equation that good teachers equal good people who are friendly, cheerful, sympathetic and moral (Getzels & Jackson, 1963). This may be true, and it may even necessarily be true,

but research has yet to demonstrate a consistent relationship between teacher characteristics and student outcome measures (Rosenshine, 1971).

If teacher training institutions are to produce effective teachers there is no doubt that teacher educators must acquire research-based knowledge as to what teacher behaviors are related to what student outcomes. But in view of the history of curriculum and pedagogical reforms and improvements, how can we assume that such knowledge can be successfully implemented through teacher training programs? Laboratory schools and demonstration teaching programs have already shown the difficulties of adopting teaching methods that require a teacher to change herself. The addition of child development and educational psychology courses into teacher preparation programs, while increasing the information given to student teachers, has not as yet given much evidence of impact on public school education. Such content has, in some instances, increased awareness of the gap between knowledge and performance, but alone has not proved to be the means for achieving the desired end.

Combs (1969) applying the concept of the "helping relationship" previously identified in a study of therapists found no significant difference in the answers of "good" and "poor" teachers--both could identify and agreed with the therapists as to what constituted the most desirable and productive relationship between students and teachers. But the "poor" teachers could not put their knowledge into action. Ginott (1972) reports from the teacher's perspective:

> What counts in education is attitudes expressed in skills. The attitudes that count are known. In fact, teachers are tired of hearing about them again and again at every conference and convention. As one teacher put it: "I already know what a child needs. I know it by heart. He needs to be accepted, respected, liked, and trusted; encouraged, supported, activated, and amused; able to explore, experiment, and achieve. Damn it! He needs too much. All I lack is Solomon's

wisdom, Freud's insight, Einstein's knowledge, and Florence Nightingale's dedication. (p. 38).

The Two Questions for Teacher Behavior Research

Two questions that must be answered by research into teacher behavior are:

1. What is the relationship between teacher behavior and student outcome?
2. Can teacher training programs produce desired change in teacher attitudes and behaviors?

These two questions take seriously the idea that teaching is a profession and that as professionals, teachers must have certain skills and attributes to give creditable performances. As with any other profession, we intuitively recognize a certain set of behaviors that mark the "born" nurse, or doctor, or artist, or salesman, or actress, or leader who had only to acquire the expertise of his field to become "professional." The "born" teacher has long been recognized on the same basis.

But the admissions policies of current teacher training programs hardly warrant the burden of "professionalism" now being placed upon their graduates. Teacher training institutions have assumed that students entering the teaching profession already possess the necessary values and attitudes that will support them in the teaching-learning situation--an assumption which is, to say the least, unfounded. The majority of the candidates for education programs are young females whose admission into teacher training programs is often viewed as entry into a socially acceptable "holding pattern" between high school graduation and marriage rather than serious candidacy for a profession. Perhaps the view of teaching as a nurturing or helping profession, coupled with the projected role of young women as potential mothers, in addition to the need for a large population reservoir

from which to fill the demand for teachers has led to this accomodation. But if improving the quality of education through its implementers is to be seriously considered and if such quality is to depend upon the use of effective teaching behaviors, we are faced with either screening for those teachers who are constitutionally capable of such behaviors or finding a way to implant effective teaching behaviors.

A Responding Educational System

The earnest efforts to improve education in the past ten years have resulted in little but cosmetic changes and semantic reforms. The majority of our school classrooms are pretty much what they were 40 years ago. Students who drop out of school and students who enter college are still saying that their education has been meaningless for the most part and has not met their "needs." We have not yet successfully responded to the consumer's complaint that education in our schools, though increasingly expensive, remains "depersonalized" and "irrelevant." When pressed for more specific information on what would be personalized or what would be relevant to their education, very few students can respond directly and pointedly. They respond instead in vagaries--not on a "knowing" level but on a "feeling" level. Educators who recognize only the cognitive process of education can dismiss such responses by questioning the intellectual capacities of complaining students. But educators who are aware of the affective significance in educational processes have recognized the legitimacy of the criticism and have looked for realistic ways to respond.

The most realistic response educators can make to the demand for relevant and personalized education is to provide preservice teachers with relevant and personalized teacher training. The curriculum reforms and

pedagogical reforms that appear blunted at the public classroom door are also dulled at the doors of many teacher training institutions. A relevant and personalized teacher training program would responsibly reflect all the cognitive objectives, the methodological skills, and the values of educational philosophy. But it would additionally include an affective and interpersonal element which would recognize the need to know and to help student teachers with their feelings as well as their fundamentals. We may not have sufficient knowledge as to what characteristics of teachers produce pupil gain, but an individual student teacher could be encouraged to find and develop her own most effective teaching behaviors with a criterion of pupil gain. To assume such an obligation will require the affective expertise described by M. M. Buchanan (1971) as "first, the ability to reach a student as a fellow human being and, second, to feed subject matter into that relationship." With such "affective expertise," educators could respond to the need for a personalized and relevant education for every student.

The Conceptual Framework of Personalized Teacher Education

The Personalized Teacher Education Program, developed by the Research and Development Center for Teacher Education at the University of Texas at Austin, has concentrated on adding the necessary affective emphasis to teacher training in order to provide both the experience of a personalized and relevant education and the means of acquiring affective expertise. As an experience, the program introduces personal interaction at successive stages of the learning process with each prospective teacher actively participating in the planning and implementation of his own training.

The term "personalization" does not simply refer to the general sense of "getting to know students better." It refers to a systematic process of assessment, feedback, and consultation for each prospective teacher and her instructors, a process specifically designed for the teacher training situation. It is through this process that the prospective teacher takes the initial steps of experiencing a personalized education and acquiring the basic self-knowledge relevant to developing affective expertise as a future teacher. It is through the repetition of this system of assessment, feedback, and consultation that the affective domain is added and integrated into the prospective teacher's program of studies.

For example, the student teacher, at the entry level, provides data about herself through self-report instruments. An assigned counselor goes over the responses and discusses them with her, as they apply to the candidate's teaching potential. The problems she anticipates or problem areas the counselor feels are indicated are discussed thoroughly. Later, the student, the counselor, and her instructors enter into consultation over her training and ability to perform in interactive situations and content areas. As the student teacher experiences her first confrontation with pupils--observing, tutoring, teaching a micro-lesson--all of these persons become involved in the processes of assessment and feedback. In this way, the affective domain becomes integrated throughout the program.

Research Base of the Conceptual Framework

The concept of a "personalized" teacher training program was developed from a complex of studies beginning at the University of Texas at Austin in the middle 1950's at what later became The Personality Research Center. Research and training procedures which had been developed at the University

of Chicago during the 1940's and 50's in a study of the relationship of psychological characteristics of business executives to effectiveness of job performance were extended and adapted to the teacher training process. In a study supported by the Hogg Foundation for Mental Health (Peck, 1958), the techniques and instruments for assessing career-relevant personal characteristics and the process of feedback were investigated as a way to train teachers. This experimental work which attempted to devise ways of improving both the self-insight and social-insight of prospective teachers was incorporated into an expanded study, the Mental Health in Teacher Education (MHTE) study, funded by the National Institute for Mental Health in 1968. The MHTE study moved into the school classroom to test the hypothesis that increased self-knowledge acquired in a supportive and constructive situation would increase the self-esteem and self-assurance of prospective teachers and would enhance their teaching performance. It was assumed that the benefits derived from the student teacher's increased understanding of the principles of mental health and deeper self- and social-insight would ultimately have a positive influence on her classroom behavior. Five instruments developed over this period of time now comprise the core of a battery of assessment instruments known as the Comprehensive Personal Assessment (COMPASS) Battery used in the Personalized Teacher Education Program: The Peck Biographical Information Form, an information gathering device as well as a projective instrument; the Bown Self-Report Inventory, a quick-scoring, self-report instrument on career-related dimensions; the Veldman-Peck One Word Sentence Completion Form, a free-response instrument scored by computer processing; the Veldman Directed Imagination Test, a projective test instructing the subject to write four brief, fictional stories about teaching, and the Adjective Self-Description, a concise and direct means of measuring major aspects of self-perception.

An important contribution to the Personalized Teacher Education program and its conceptual framework was made through the findings of the MHTE study. These findings indicated that problems of prospective teachers coincided frequently in the areas of:

Attitudes toward authority (e.g., relationships with cooperating teachers, supervisors, principals, instructors),

Attitudes toward children (e.g., grade level preferences, positive-negative responses to individual children, selective preferences for boys or girls),

Attitudes toward work (e.g., content adequacy, preparation for teaching, amount of participation, attendance), and

Attitudes toward teaching (e.g., continuation in the program, persistence in teaching).

Identification of these potential problem areas gave direction for determining and defining the goals of the Personalized Teacher Education Program. With additional investigation and analysis of these problem areas after the initial work of the MHTE project, three domains of competence were designated in the conceptual framework: (I) Intrapersonal Competence; (II) Interpersonal Competence; and (III) Career-related Competence. Each one of these three competency domains corresponds to problem areas which appear during the teacher training sequence, and individual goals, appropriate to these three areas, are set for individual student teachers.

It is interesting to note that during roughly the same period of time that the Personalized Teacher Education Program was being developed, Combs (1969) employed a different approach to arrive at similar domains of competence needed for effective teachers. Having defined teaching as a "helping" relationship, Combs examined the perceptual differences between those persons in helping professions (e.g., counselors, nurses, priests) who were deemed "effective" in their roles and those who were considered "ineffective." Four categories of differences were noted:

I. General perceptual organization. Is he more interested in people or things? Does he look at people from the outside, or does he try to see the world as they see it? Does he look for the reasons people behave as they do here and now, or does he try to find historical reasons for behavior?

II. Perceptions of other people. Does he see people generally as able to do things or unable? As friendly or unfriendly, as worthy or unworthy? As dependable or undependable?

III. Perceptions of self. Does he see himself as with people or apart from them? As able or unable? As dependable or undependable? As worthy or unworthy? As wanted or unwanted?

IV. Perceptions of the professional task. Does he see his job as one of freeing people or controlling them? Does he see his role as one of revealing or concealing? As being involved or uninvolved? As encouraging process or achieving goals?

In helping a student teacher achieve more competence in self-knowledge, in ability to know and interact with others, and in achieving those interactive skills and values which relate to her profession, the Personalized Teacher Education Program helps student teachers toward those perceptions indicative of more "effective" helpers.

In the following table, examples of individual program goals for a student in the Personalized Teacher Education Program are shown as they relate to the three Domains of Competence.

Table 1-1. Domains of Competence.

Intrapersonal	Interpersonal	Career-Related
Personal goal achievement	Awareness of others	Use of teaching resources
Self-confidence	Ability to relate to others	Classroom management
Independence	Responsiveness	Knowledge of subject
Realistic self-perception	Appropriate empathy	Knowledge of child development
Congruence (a match between feeling and behavior)	Receptivity to feedback	Pupil evaluation skills
	Supportive, positive, and encouraging toward pupils	Alternative teaching styles

The prospective teacher works toward these goals within the supportive relationships established in her program to which can be added, at appropriate levels of application, the content knowledge and experiences necessary to achieve affective expertise as well as professional competence.

The Concerns Theory

It is essential to the basic concept of "Personalization" that the difference between "concerns" and "needs" be understood. The "concerns" of the individual are subjective in nature, incorporating perceptions, values and attitudes. The "needs" of an individual may or may not be subjective, and actually, may not be concerns. Concerns relate to those needs reported by the individual, and not the "needs" ascertained by objective assessment. Teacher education programs have long objectively ascertained the needs of prospective teachers and have generally restricted attention to needs in the cognitive areas. Only recently have teacher educators considered the relationships between the needs and concerns of student teachers and the training program modifications suggested by such relationships.

The Concerns Theory as initially stated by F. F. Fuller (1969a) grew out of the analysis of recorded typescripts of student teaching seminars and interviews with student teachers. These records, over an extended period of time, were used in the identification and classification of problems which student teachers experienced and the concerns they expressed about these problems. These expressed concerns when grouped into definable developmental and sequential stages showed that the early concerns of student teachers were characterized by a concern for self and self-protection, while the later concerns of student teachers and inservice teachers satisfied with their teaching were characterized by a concern for others, for relationships with others and for pupil learning. The identification of the concerns of student teachers and the sequential nature of these concerns was undertaken in the Personality Teacher Education and Teaching Behavior (PEB) study funded by the U. S. Office of Education (Fuller, Peck, et al., 1969).

Stated in its simplest terms, the Concerns Theory conceptualizes the learning process for a prospective teacher as a natural flow from concerns for Self (trainee) to Task (teaching) to Impact (pupil). Since learning in this sequence proceeds from the self, the prospective teacher must be the starting point for planning and structuring any learning experience.

The physical, mental and emotional state of the prospective teacher plays an important role in the shift of focus from self to task to impact. Any cognitive or affective impediment results in a slower, more labored shift of focus to task and impact that can, in turn, result in a failure on the part of the prospective teacher to obtain minimal teaching competencies from the training program.

One function of measuring the concerns of the preservice teacher is to identify these learning impediments. The initial concern of each trainee as she

is confronted by a new learning experience will be for the self. This subjective assessment of the learning situation yields concerns which can give the teacher educator access to the motivations and perceptions of the learner and an entry point for the development of the trainee. The teacher educator uses concerns as a basis for structuring affective and cognitive experiences which can shift the trainee from concern for self to concern for task and, ultimately, to a concern for the impact she is having upon pupils.

Initial concerns for self, which include concerns for self-protection, must be reduced in order to focus the trainee upon the teaching task, i.e., learning objectives and the teaching environment. The first step in the training of prospective teachers, then, should be knowledge of self or intrapersonal knowledge. This knowledge can be gained through the assessment and analysis of data obtained from self-reports, reports by peers and supervisors and behavoral observations made in a systematic fashion. By measuring the concerns of the preservice teacher, it is possible to help the student become aware of her self-concerns and then to clarify her motivations, often conflicting ones, in an effort to identify personal goals, to resolve anxiety, and to reduce defensiveness, particularly when dealing with learning goals and objectives that are not consonant with the value system of the student and, therefore, not perceived as desirable goals. This process is important in helping the student perceive herself as capable of achieving goals which she never considered possible, didn't know existed or know about but considered incompatible with her own value system.

For example, consider an assignment in a social science methods course in which each student teacher is to teach a mini-lesson. For student "A," the assignment immediately invokes a concern for self because she cannot comfortably stand before a class of her peers. She prepares her material and lesson plan well but, because of her apprehension, gives a poor performance.

Student "B" in the same class is delighted with the assignment. She is happy to appear before the group and enjoys a chance to perform. She does not prepare the content of her lesson carefully, however, so she too gives a poor performance, enthusiastic but contentless and uninteresting to her students.

If the concerns of student "A" are known, one purpose of the mini-lesson would be to help her acquire ease before her peer group. Having failed too, student "B" may now have a concern for preparation that can be used as the focus of her next assignment.

There must be, however, a linking event in this process, an awareness of the relationship between the student's present status and goals as well as an awareness of previously unperceived goals which are possible for her. This linkage is achieved through assessment and feedback of many aspects of her experience, including concerns, as a part of each learning experience. Through the judicious structuring of learning experiences based upon level of concern, the preservice teacher is led to an awareness of the task and aroused toward achieving the learning objective as a personal and relevant learning experience. The progression of concerns from self to task to impact can be reflected both in the macrocosm of a total training program and in the microcosm of a single training experience.

The Personalized Teacher Education Program

Both the conceptual framework and the Concerns Theory are implemented in the Personalized Teacher Education (PTE) Program. The goal of the PTE Program is to help each prospective teacher develop his own effective teaching behavior by achieving his individual goals in the three domains of competence,

i.e., Intrapersonal, Interpersonal, and Career-Related. There are different goals for different students, variant treatments for diverse personalities, and different growth rates in different directions. It cannot be assumed that the PTE Program will be good for each student in the same way, nor will one ideal "teacher" type emerge. This is in keeping with the "self as instrument" concept which has defined teaching effectiveness as "a function of how the teacher combines his knowledge and understanding with his own unique way of using self to be helpful to others" (Dinkmeyer, 1971, p. 617).

The PTE Program is administered through "blocks," administrative groupings of professional courses. Blocks may vary from 9 to 18 credit hours. Students registered in a block will have the same courses at the same time under the same instructors for one full semester. PTE blocks are staffed by teams consisting of all course instructors and one or more counseling psychologists, depending on the size of the block. The cooperating classroom teachers are added to the team during the Observation and Student Teaching course work. Team planning begins prior to the school semester to assure an interdisciplinary approach, to ensure the best use of all learning experiences, and to plan opportunities for feedback. The counselors help by supplying suggestions and insights on student problems and assist in coordinating work done in the public schools.

The following describes the specific activities which comprise the observation and student teaching semesters of a Personalized Teacher Education Program.

Activities of the Observation Semester

1. One week in public school as teacher aid.

   Prospective teachers get a realistic picture of the teaching role, behavior of students in classrooms, functions of teachers and the school environment.

2. Prospective teachers attend orientation meeting of their block, take the COMPASS Battery of tests.

   Prospective teachers get acquainted with faculty members and their fellow prospective teachers and learn basic concepts of personalized education. Slide-tapes such as "Meet Your Cooperating Teacher" (Fuller and Newlove, 1969) may be used.

3. Personal Assessment Feedback.

   Prospective teachers attend counseling session with assigned counselor to receive feedback from COMPASS Battery and open channels of communication for discussing problems and personal goals.

4. Instructor conference schedule.

   Prospective teachers have time with each instructor to become acquainted and have their perception of their first in-school experiences assessed by the instructors. Instructional modules may be used to help develop non-instructional classroom skills.

5. Campus classwork.

   Regular course work in methods courses, educational psychology, etc., occur either in space provided at the public school or on campus. Prospective teachers are required to integrate learning with classroom experiences.

6. Observation.

   Prospective teachers tutor small groups or individual students and help in lesson preparation. Prospective teachers visit alternative schools to familiarize themselves with other settings and socio-economic conditions.

7. Videotaping of short lesson.

   Prospective teachers prepare and teach a short lesson which is videotaped.

8. Videotape feedback.

   Instructors and counselors provide prospective teachers focused feedback on videotaped lessons.

9. Seminars with invited speakers, films.

   Prospective teachers receive additional views of teaching as a profession, learn audio-visual forms, preparation of material, use of machines, and community resources.

10. Prospective teachers prepare and present one regular teaching session in content major.

    Prospective teachers are videotaped and comparisons are made with the earlier videotaped performance. (See 7 above.)

Activities of the Student Teaching Semester

1. Prospective teachers are assigned to a school and classroom different from observation site.

2. Prospective teachers attend seminars relating to subject matter, the analysis of teaching behavior, and individual problems.

3. Prospective teachers teach whole class and small groups. Conferences between college supervisor, classroom supervisor, counselor and prospective teacher take place.

4. Prospective teachers assume full teaching role for experience in classroom instruction and management. Evaluation and consultation with classroom supervisors and counselor take place.

5. Prospective teachers continue seminars and course work on campus.

As with all other experimental teacher training programs, a Personalized Teacher Education Program operates within certain real-world constraints. The students participating in the program must have the course content and experiences required for graduation and certification by the College of Education and the State. In addition, the required public school experiences can only occur through the joint cooperation of the College of Education and the School District, each of whom sets and maintains its own standards.

The PTE Program, while adding the affective dimension to the training process, does not necessarily alter course content, though it may alter its sequence. Modules in several content areas have been developed to allow for individualized course work, through self-pacing. Generally, the PTE Program differs from a traditional program by demanding more time and

flexibility from prospective teachers, staff, and cooperating institutions and by increasing the scope of educational training to include affective experiences. To foster the attainment of this latter objective, a PTE Program measures the incoming personality and attitudinal characteristics of the trainee and plans training experiences that match the trainee's current level of affective functioning. Trainees thereby receive different sets of instructional experiences and are expected to achieve both affective and cognitive outcomes to differing degrees depending upon their entering personality and attitudes.

The remainder of this report is devoted to a further explication and an evaluation of the personalized model of teacher training. The following chapters report the extent to which a personalized program as it was implemented in a small-scale field try-out fostered the professional growth of its trainees. Underlying the evaluation design employed for this study was the belief, implicit in the personalized model, that any one training program or single set of training experiences may not be best for all students and that any study of teacher training should seek to identify those students for which a particular program is best suited. Therefore, the following study was designed to assess the extent to which the effectiveness of the personalized and traditional models of teacher training is a function of the entering personality and attitudinal characteristics of the trainee.

SECTION II.

RESEARCH REPORT

Chapter 2

OVERVIEW OF THE EVALUATION DESIGN

Recent research on teacher behavior (Rosenshine, 1971) can be categorized into three general areas: studies about the intrapersonal behaviors of teachers, i.e., their personalities and attitudes; studies about the interpersonal behaviors of teachers, i.e., their interactive modes with pupils; and studies about the subject-matter competence of teachers, i.e., behaviors related to the content they teach.

A major goal of most teacher training programs is to teach the interpersonal behaviors and subject-matter competencies most frequently needed in the act of teaching and which relate most directly to the affective and cognitive growth of the school child. These training programs strive to achieve this goal through course instruction, classroom observation and student teaching experiences that focus upon the attainment and utilization of subject matter competence. Due to conventional commitment to these ends, specific personality and attitudinal traits of the prospective teacher, while often hoped-for results of these experiences, usually are considered indirect outcomes of or spin-offs from the planned instructional sequence. The conventional model of teacher training as it is customarily applied at teacher training institutions contains four distinct characteristics: (1) general university course work in the field of education, in the sciences, social sciences and the humanities; (2) content-related methods courses from which the prospective elementary school teacher learns how to teach mathematics, science, social science and language arts and the

prospective secondary school teacher learns how to teach either one or two of the above or a more specialized discipline; (3) a planned sequence of classroom observation wherein the trainee observes and sometimes records teaching behaviors as they occur in actual classroom settings; and (4) student teaching wherein the trainee learns to apply his skills in classroom management and group instruction under the supervision of an inservice teacher.

The conventional model of training is characterized by fixed program goals attained through a fixed instructional sequence. It places initial emphasis upon the accumulation of knowledge about subject-matter content and teaching methods with later emphasis upon the application of knowledge and methods in an ongoing, teaching-learning environment. Training programs that focus upon specific competencies during the training sequence and that require of the trainee minimal levels of attainment for these competencies often are referred to as performance- or competency-based.

Peck (1972a) has suggested that the appropriate utilization of content-related behavior may be dependent on both the intrapersonal and interpersonal behavior of the trainee and to the extent that intrapersonal and interpersonal growth is limited, the acquisition of effective content-related behaviors may be more difficult or even unattainable. This perspective suggests that a trainee who lacks self-confidence, for example, will experience difficulty in becoming an effective teacher even though he may be capable of attaining a high level of subject-matter competence. A teacher training program which focuses on intrapersonal and interpersonal behavior as well as subject matter competence employs what will be referred to in this report as the personalized approach or model. Such programs differ from conventional training in that they focus upon the development of intrapersonal and inter-

personal behavior in conjunction with the attainment of subject-matter competence.* Examples of the intrapersonal, interpersonal and content-related behaviors that a personalized program might seek to develop were noted in Table 1-1.

The personalized alternative to the conventional model of teacher training is one that may include all of the components and program goals of the conventional or competency-based model but that, in addition, takes into consideration the affective development of the trainee. Rather than leave the affective growth of the trainee as an indirect function of a broader instructional strategy, this model, as a part of the training program, adjusts or alters the instructional sequence to include additional experiences that foster the personality and attitudinal characteristics of the trainee that are thought to be prerequisites to effective teaching. This model assesses the affective development of the trainee, feeds this assessment data back to the trainee in a counseling session and plans professional experiences for the trainee that foster his affective development in ways both he and the teacher educator deem most appropriate.

Consider, for example, a prospective teacher who upon entering a training program is given a battery of instruments designed to measure attitudinal and personality traits related to effective teaching. Moreover, let us suppose that on an anxiety measure the trainee scores two standard deviations above a reference group comprised of all preservice

---

*Some teacher training institutions using the personalized model include Brigham Young University, Georgia State University, Kansas State Teachers College, Northern Illinois University, The University of Alabama, University of Colorado, The University of Houston, The University of Texas and Western Kentucky University.

teachers who have thus far entered the program. The procedure employed in the personalized model includes feeding back this information to the trainee in addition to other data confirming whatever strong points he may have, e.g., warmth for children, enthusiasm for the training program and dedication to teaching. The personalized program then plans an instructional sequence based upon the trainee's current level of affective development. In this instance, the personalized program might plan intermediate experiences that introduce the trainee to teaching in a nonthreatening, less anxiety-evoking setting than might be appropriate for his peers. Videotaped performances without the presence of peers, more frequent or earlier experiences with school children in small groups and consistent consultation with the teacher educator might be in order before the trainee is asked to perform the more routine cognitive sequence of instruction.

General differences in the conventional and personalized models of teacher training may be noted in Figure 2-1. These, however, are stereotypic versions as some programs espousing either model may differ both in degree and in kind.

Which type of training program--conventional or personalized--is more effective? Attempts to answer this question in general may not be productive in that any one approach may not be more effective than another for every prospective teacher. It is not unlikely that one prospective teacher may profit more from a conventional program and another from some other, more specialized program. Personal traits and training programs may interact thereby suggesting that no one type of training experience may be best for every student. Such trait-program interactions suggest that a prospective teacher should be assigned, when it is feasible, to that training experience that is likely to be most effective for him.

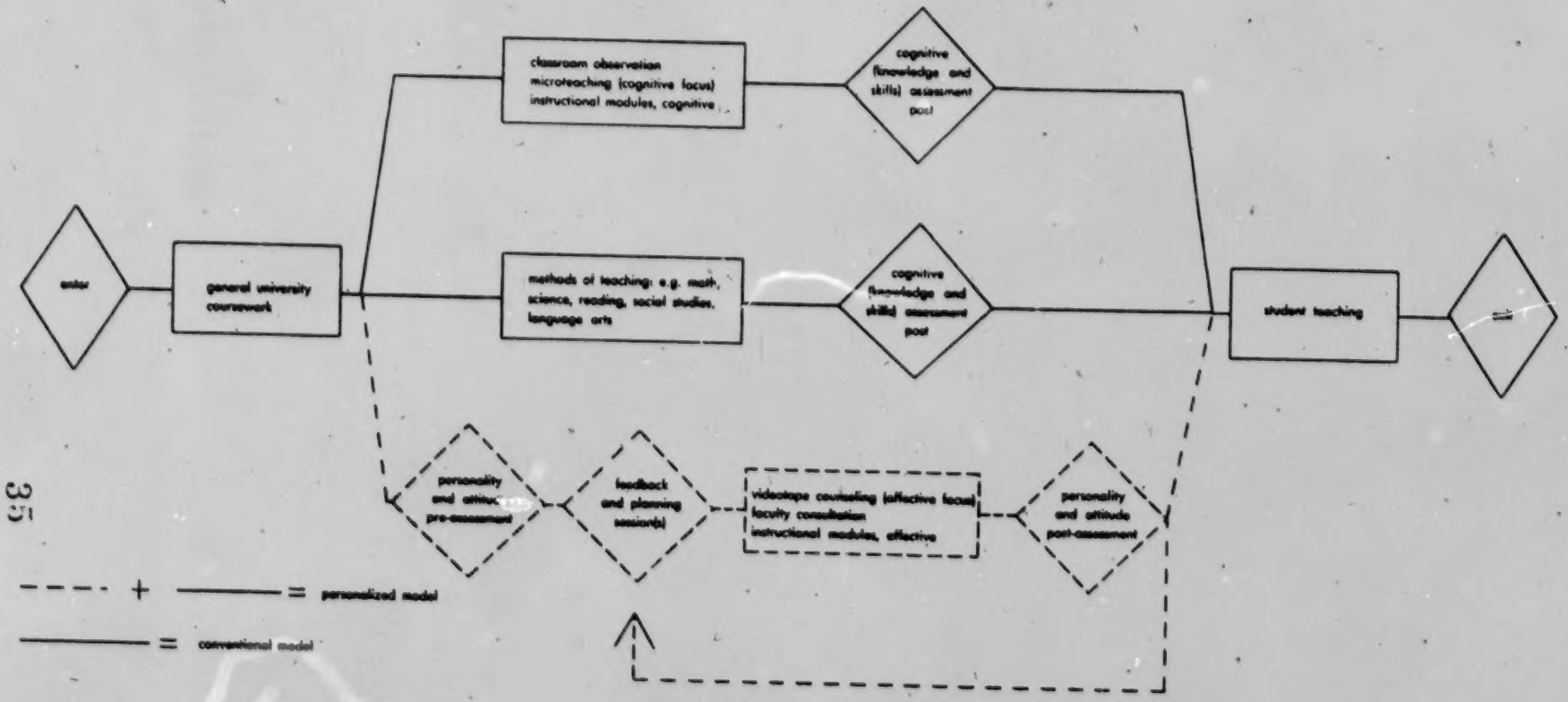


Figure 2-1. General Differences Between the Conventional and Personalized Models of Teacher Training.

Objectives. The objectives of the present research were (1) to examine the effects of the traditional and personalized approaches to teacher training as they affect teaching behavior and pupil evaluations of teaching and (2) to identify interactions between training approaches on the one hand and the personality and attitudes of the prospective teacher on the other.

Sample. Seventy-seven teacher trainees at The University of Texas were selected to participate in the evaluation study. Thirty-nine of these students voluntarily participated in a personalized teacher education (PTE) program, while the remaining 38 students voluntarily participated in a traditional program (non-PTE). All but two of the trainees were females. Trainees were assigned to student teaching at one of four public elementary schools in Austin, Texas, in the following manner:

School 1:

20 student teachers who received the first semester of the PTE program in Spring 1972 and who received the second semester of the PTE program in Spring 1973. (Experimental Group)

School 2:

22 student teachers who had received no PTE training but who had gone through the University of Texas (UT) conventional program concurrently with the experimental students in schools 3 and 4 below. (Control Group)

School 3:

10 student teachers who received the first semester of the PTE program in Fall 1972 and who completed the second semester of the PTE program in Spring 1973. (Experimental Group)

7 student teachers who had received no PTE training but had gone through the UT conventional program concurrently with the experimental students in schools 3 and 4. (Control Group)

School 4:

9 student teachers who received the first semester of the PTE program in Fall 1972 and who completed the second semester of the PTE program in Spring 1973. (Experimental Group)

9 student teachers who had received no PTE training but had gone through the UT conventional program concurrently with the experimental students in schools 3 and 4. (Control Group)

Treatments. Both the traditional and personalized programs in the present study included university course work, classroom observation and practice teaching. The unique components of the personalized program were (a) repeated counseling sessions with the prospective teacher for engendering attitudinal and personality characteristics related to effective teaching, (b) the differential assignment of instructional tasks and activities based upon the specific attitudinal and personality characteristics of the prospective teacher, (c) self-observation of teaching behavior through videotaping, and (d) affective feedback and assignment of tasks and activities related to this self-observation.

General hypotheses. Several hypotheses are implicit in the design of the Personalized Teacher Education Program.

(1) Because the thrust of a personalized program is to provide for individual differences by varying the rate and kind of learning for each student, interactions between student traits and training programs, not main effects between programs, should be found with personalized approaches to training. Analysis of mean differences between training programs should reveal nonsignificant differences between the personalized and traditional models of teacher training.

(2) When entering levels of personality traits and attitudes are considered, a personalized appraoch will be found to be more effective in engendering teaching behaviors than a traditional approach for some personality traits and some attitudes.

(3) Specifically, a personalized program will be more effective in fostering the individual growth of students who score below average on personality and attitudinal traits related to teaching and thereby are in most need of a personalized treatment than students who score above average on these traits.

Methods. Personality and attitude scales were administered to students in the conventional and personalized programs upon entry into training, and these measures provided the trait variables investigated in the present study. Personality, attitude and teaching effectiveness measures were chosen to cover a broad range of behaviors consistent with the three domains of competence, i.e., intrapersonal, interpersonal and career-related, posited in the conceptual framework of the Personalized Teacher Education Program. These behaviors were measured via self-reports, reports of others and systematic observational coding systems as noted in Table 2-1. Three direct observational coding systems were used to collect the observational measures of teaching behavior--the Classroom Observation Scales (Emmer and Peck, 1973), the Instrument for the Analysis of Science Teaching (Hall, 1969), and the Fuller Affective Interaction Record (1969b). Teaching effectiveness (criterion) variables were measured at the end of the practice teaching semester, the last semester in the training sequence. The instruments used to measure these variables are discussed in the following chapter.

A trait-treatment interaction design was employed to determine differences between the personalized and conventional programs for different personalities and attitudes. The personality and attitude measures served as the trait variables, program (conventional vs. personalized) served as the treatment variable and the measures of teaching effectiveness served as the criterion variables. Trait-treatment interaction methodology

Table 2-1.
Variables Measured Classified According to the Type of Measurement by Domain of Competence Matrix.

Type of Measurement	Domain of Competence: Intrapersonal	Interpersonal	Career-Related
Self-Report	ASD: Efficiency Anxiety Idealism Attractiveness SRI: Self Reality Hope Readiness Assessment: Self-Concern Perceptive About Self Teacher Concerns: Personal Concerns Teacher Beliefs: Personal Adjustment Ideology	ASD: Attitude Behavior Introversion SRI: Children Others Authority Parents Readiness Assessment: Concern for Children Perceptive About Children's Behavior Teacher Concerns: Concern for Pupils Teacher Beliefs: Consideration of Student Viewpoint	SRI: Work Readiness Assessment: Concern for Impact Motivated to Teach Teacher Concerns: Professional Concerns Professional Plans and Affiliations: Job-Seeking Behavior Motivation to Teach Profile of Learning Priorities: Competent Management Professionalism Flexibility Responsibility Teacher Beliefs: Student Autonomy vs. Teacher Control
Other-Report	Readiness Assessment (College Super.): Perceptive About Self Self-Concern Readiness Assessment (Pub. School Super.): Perceptive About Self Self-Concern SET 2: Unreasonable Negativity	Readiness Assessment (College Super.): Concern for Children Perceptive About Children's Behavior Readiness Assessment (Pub. School Super.): Concern for Children Perceptive About Children's Behavior SET 2: Rapport Fostering of Self-Esteem	Readiness Assessment (College Super.): Concern for Impact Motivated to Teach Readiness Assessment (Pub. School Super.): Concern for Impact Motivated to Teach
Observation		IASTV2: Teacher Praises Affective Response--Positive FAIR: Teacher Nurtures COS: Positive Affect Negative Affect	IASTV2: Question--Open Question--Closed Accepts Student Statements--Restates Accepts Student Statements--Questions Accepts Student Statements--Short, Non-evaluative Confirmation Direction--Managerial Direction--Procedural Lecturing, Giving New Information Reading Aloud to Class Controlled Silence Controlled Silence--Preparing Material Student Statements--Closed Student Statements--Open Student Statements--Reading Aloud Student Questions--Substantive Closed Student Questions--Substantive Open Student Questions--Procedural Closed Student Activity--Overt Student Activity--Covert Student Activity--Group Overt Student Activity--Class/Group Verbal Division of Student-to-Student Interaction Non-functional Behavior Teacher-Talk to Student-Talk Ratio Extended Teacher-Talk to Extended Student-Talk Ratio Flexibility Ratio FAIR: Teacher O.K. Teacher Helves Teacher Confirms Teacher Ponders Teacher Corrects Teacher Tangential Teacher Initiates Teacher Manages Teacher Lectures Student Zeal Student Explores Student Usual Student Questions Student Suggests Student Rejoices (for Self) Student Admits Student How! Student Brings Out Teacher Solitary Work Student Solitary Work COS: Level of Attention Teacher-Initiated Problem-Solving Pupil-to-Pupil Interaction Teacher Presentation Higher Cognitive Level Student Behavior Passive Pupil Behavior Convergent Evaluative Interaction Task Orientation Clarity Enthusiasm Career-Related Behavior

differs from traditional factorial designs in that trait variables commonly dichotomized or trichotomized to fit the factorial structure of analysis of variance are not divided into discrete categories but rather are used in their continuous form to describe as many different types of trainees as there are observed values of a particular trait.

The general methodology may be summarized in three steps. The first step is that of correlating entering personality and attitudinal traits with outcomes for each program. If, for example, the trait-criterion correlation is positive for one program and negative for a second, the first program is likely to be more effective for individuals scoring high on the trait; the second program, for individuals scoring low on the trait. A second step is to calculate trait-criterion within group regression slopes and the extent to which the regression slopes differ, i.e., are heterogeneous, across programs (Edwards, 1968). Should regression slopes significantly differ, a third step is employed to determine the exact regions of trait values for which the programs are significantly different (Walker and Lev, 1953; Borich, 1971). Figure 2-2 pictures a hypothetical study for which there are significantly different regression slopes for two programs with regions of significance to the left and right of the point at which regression lines intersect. Students with trait values above point B should be assigned to Program I while students who score below point A should be assigned to Program II. For students scoring between points A and B, both treatments are equally suitable for producing the criterion behavior and such individuals should be assigned to the least costly program. Our discussion now turns to the specific instrumentation for this study.

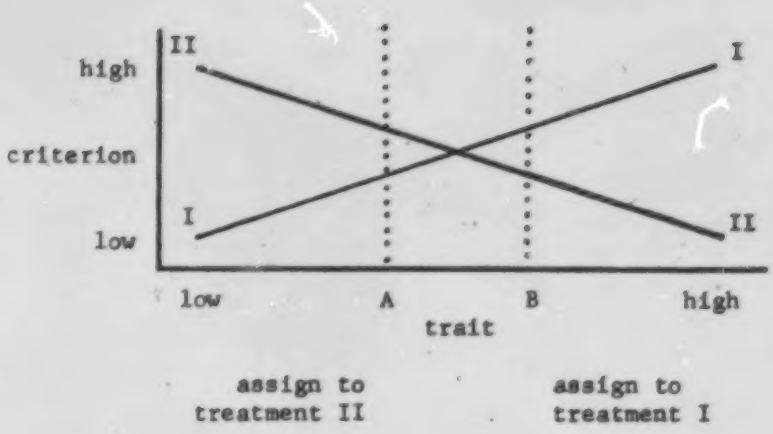


Figure 2-2. A Hypothetical Trait-Treatment Interaction.

Chapter 3

INSTRUMENTATION

Instruments used in the present investigation measured two types of variables: trait variables and criterion variables. Several additional instruments were administered, but not analyzed. These included procedural forms (such as for subject identification) as well as instruments which yielded data not analyzed due to time constraints and/or the conviction that the data gathered at one or more of the sites was invalid. The variables which were analyzed in the present investigation have been classified according to the Type of Measurement X Domain of Competence Matrix. This classification appeared in Table 2-1.

The discussion below is divided into sections based on the classification of instruments by type of variable measured. All instruments which were used, but for which no data analysis is reported in this document, are briefly described under the final subheading, "Additional Instruments."

Information bearing on the reliability and validity of each instrument is reported whenever it is available.

Criterion Variables

The majority of the criterion variables for the present study were based on the two videotaped lessons taught by student teachers. Accordingly, observational coding variables will be discussed first. Descriptions of the remaining instruments will follow in alphabetical order.

An initial videotaping (pretest) was made after the student teacher had several weeks of exposure to her pupils, and a final taping (posttest) was recorded as late in the semester as possible. Approximately eight

weeks of actual classroom experience intervened between the two tapings. This was the maximum interval possible, given the constraints of the University of Texas and Austin Independent School System calendars. The procedures used for the videotaped lessons are described below under "Data Collection Procedures," and the videotape guidelines distributed to student teachers are reproduced in Appendix B.

Videotapes of the student teachers' lessons were coded according to three different systems: Instrument for the Analysis of Science Teaching, Version 2 (IASTV2); the Emmer-Peck Classroom Observation Scales (COS); and the Fuller Affective Interaction Record (FAIR). All three observational systems quantify behaviors which occur in the classroom situation. Published manuals are available for the IASTV2 (Hall, 1972), the COS (Emmer, 1971), and the FAIR (Fuller, 1969b) systems.

Two trained coders independently scored each videotape using each of the three systems in turn. Data analyzed for the present investigation consist of mean scores across the two coders for each variable.

The three observational systems differ with regard to both the behavioral categories employed and the observation interval. The observation interval refers to the fixed length of time for which the coder views the videotape prior to recording what behaviors occurred. The IASTV2 has a very short observation interval of approximately 3 seconds, while the COS employs a 15-minute interval. The FAIR system uses a continuous coding process which can be set to pulse at any rate between one and five times per second. Each time it pulses, it records a repetition of the previously observed behavior until the coder punches a different key to indicate behavior change.

Instrument for the Analysis of Science Teaching, Version 2 (IASTV2)

The 32 behavioral variables provided by this instrument are listed and described below. Each of these variables corresponded to a particular category of behavior. The score for a variable was the relative frequency of occurrence of the behavioral category in question--specifically, 100 times the frequency of occurrence for that category divided by the frequency of occurrence for all categories. The score for a variable is thus interpretable as the percentage of time the behavior in question occurred. The reliability coefficients reported for each variable are intraclass correlations (Winer, 1962; Medley and Mitzel, 1963) obtained in category-by-category analyses on two sets of videotapes made by preservice elementary teachers, as reported in Hall (1972). The intraclass correlation is obtained from a comparison involving the amount of variation among observers and the amount of variation among classes or teachers. If there is as much variation among observers (coders), averaged across observarions, as among classes or teachers, then there is no basis for assuming that observers agree beyond what would be expected by chance. High coefficients indicate that most of the variability among scores is caused by differences among teachers or classes, rather than disagreement among coders.

It should be pointed out that in at least some cases low reliability coefficients may be attributable to a low frequency of occurrence for a behavior rather than to coder error.

Data analyzed for the present study were mean scores averaged across two coders. Reliability coefficients, as reported by Hall (1972), appear in parentheses following the variable name.

1. Acceptance of feelings (empathy); by teacher. (0 and .16) This category includes joking, when not at the student's expense. It also includes such behaviors as, when it is said sincerely, "I know that this is difficult, but let's try it anyway." If the teacher touches or puts her hand on the shoulder or head of a child, as a form of encouragement, then this behavior would be scored as acceptance of feelings.

2. Praise; by teacher. (.34 and .53) This category involves praise--"That's a good job, John." However, this would not include the use of "good" as a response to every student's statement. This would be a verbal habit and therefore would no longer have any meaning for a student.

3. Acceptance of student statements-restates; by teacher. (.89 and 94) The teacher restates, giving a limited expansion or clarification of student statements.

4. Acceptance of student statements--questions; by teacher. (.19 and 0) The teacher questions student statements, inducing the students to clarify or expand their statements.

5. Acceptance of student statement--short, non-evaluative confirmation; by teacher. (.38 and .76) The teacher responds to students' statements with non-evaluative confirmations such as "yes," or "okay," where no value judgment is implied.

6. Questions--closed; by teacher. (.83 and .92) The teacher asks narrow, specific, or channeled questions which require a specific response. Such questions require application of simple or complex skills to produce a convergent, or memorative, response.

7. Questions--open; by teacher. (0 and .55) The teacher asks broad questions which provide opportunities for students to be original in their

responses. The teacher asks questions or makes evaluative statements which evoke further thought.

8. Direction--procedural; by teacher. (.41 and .89) The teacher gives directions and procedures for substantive behaviors, where an immediate student response or behavior is required.

9. Direction--managerial; by teacher. (.35 and .91) The teacher gives directions which do not deal directly with lesson content. Immediate behavior is required from students.

10. Give substantive information--lecturing. (.88 and .93) The teacher lectures, providing substantive facts or calculations. This includes writing new information on the chalkboard.

11. Give substantive information--previous information; by teacher. (0 and 0) The teacher repeats or reviews information presented during a previous class period. Reported reliability coefficients are both 0. However, as Hall (1972) points out, this and some other low reliability coefficients on the IASTV2 may be due to the low frequency of the behaviors rather than to coder error.

12. Give substantive information--reading aloud; by teacher. (0 and .93) The teacher reads aloud from a textbook, teacher's commentary, or other source.

13. Justification of authority; by teacher. (.51 and .71) The teacher engages in unconstructive criticism or rejection of student ideas or behaviors, showing unmistakeable displeasure. This category of behavior includes self-justification and disciplinary statements of a critical or defensive nature which rely upon teacher's position of authority and have negative

affective mannerisms involved.

14. Controlled silence--demonstration; by teacher. (.73 and.33) The teacher controls the class, causing students to remain silent while he performs a demonstration before the class. No verbal behavior takes place during these intervals.

15. Controlled silence--controlled silence; by teacher. (0 and .12) The teacher has the initiation of action within her control, as in the period following a teacher question before the teacher names a specific student to answer, or after a strong reprimand.

16. Controlled silence--looking at notes; by teacher. (0 and .79) The teacher reviews her notes, lesson plans or other materials while the class remains silent and waiting.

17. Controlled silence--preparing material; by teacher. (.18 and .59) The teacher prepares, adjusts, or distributes instructional apparatus, equipment or manipulative materials. This includes, for example, handing out or gathering in papers, and arranging a visual device.

18. Student statements--closed; by student. (.85 and .85) Students' statements reflect memorative, or convergent, thought. The statement does not reflect originality in thinking and may occur in response to a closed teacher question.

19. Student statements--open; by student. (.15 and .36) Students make statements which reflect evaluative, or divergent, thinking. Some original student ideas not previously discussed in a class period are included in this coding.

20. Student statements--reading aloud; by student. (0 and 0) Students read from textbooks, papers, the chalkboard, or other references.

21. Student question--substantive closed. (.71 and .93) Students ask substantive questions (related to the substantive issues of the lesson) which are convergent and memorative in nature.

22. Student questions--substantive open. (0 and 0) Students ask questions related to the substantive issues of the lesson which are divergent, or evaluative, in nature.

23. Student questions--procedural closed. (.70 and .60) Procedural questions (related to the methodology of doing the lesson or to matters such as permission to leave the room) are raised by students.

24. Student questions--procedural open. (0 and .71) Students ask procedural questions which can also be characterized as divergent or evaluative in nature.

25. Affective response--positive; by student or teacher. (0 and .76) Expressions of enthusiasm, joy, anticipation, pleasure, approval, or excitement are emitted by either the teacher or the students.

26. Affective responses--negative; by student or teacher. (0 and 0) Expressions of disappointment, negative attitude or reaction by students or teacher are coded here. This category includes evidence of hostile feelings, resentment, sarcasm, or directed anger, all of which must be judged more on the basis of mood, intonation, and intent than on what is actually said.

27. Student activity--overt. (.20 and .91) Students engage in purposive activity such as activity where students are working individually and manipulating materials, or students walking to the chalkboard and writing would also be coded in this category.

28. Student activity--covert. (0 and .52) Students engage in purposive

but silent, internal behavior such as reading silently or thinking.

29. Student activity--group overt. (.88 and 0) Students are actively engaged as in Category 27, but this activity is performed in groups rather than individually.

30. Student activity--class/group verbal. (.87 and .82) The class acts in unison, giving a verbal response. All students need not be giving the same response.

31. Division of student-to-student interaction. (.71 and .77) This coding records an exchange between students without the teacher as mediator, but with the attention of the class and the teacher.

32. Non-functional behavior. (0 and .64) Undirected, purposeless behavior takes place. No instruction is taking place. This generally occurs in periods immediately following a class, when the teacher obviously has no control over student activity.

Two ratios based upon IASTV2 categories are also considered. These ratios are intended to provide an indication of the over-all character of teacher-pupil interaction in the classroom.

33. Teacher-talk to student-talk ratio. (.92 and .78) This ratio was calculated by summing the scores for variables 1 through 13; this sum was then divided by the sum of the scores for variables 18 through 24.

34. Extended teacher-talk to extended student-talk ratio. (.73 and .68) This ratio was calculated in the following manner. First, the frequency of occurrence of the same behavioral category in two successive observation intervals was determined for the different behavioral categories. Such frequencies can be thought of as frequencies of extended occurrence. Second, the extended teacher-talk to extended student-talk ratio was then calculated

as the sum of the frequencies of extended occurrence for variables 1 through 13 divided by the sum of the frequencies of extended occurrence for variables 18 through 24.

Fuller Affective Interaction Records (FAIR).

The 29 behavioral variables provided by this instrument are listed and described below. Each of these variables corresponds to a category of behavior. The score for a variable was the relative frequency of occurrence of the behavioral category in question--i.e., 100 times the frequency of occurrence for that category divided by the frequency of occurrence for all categories. The score for a variable is thus interpretable as the percentage of time the behavior in question occurred. The reliability coefficients reported here are taken from Fuller (1969b), and represent the interjudge consensus obtained between two trained coders on a sample of 34 videotapes.

1. Teacher Values. (.58) The teacher appears to recognize and value feelings. He identifies, shares, listens attentively, or gives unqualified acceptance.

2. Teacher Nurtures. (.90) The teacher gives focused encouragement. He guides, hints, or gives praise or approval to a previous behavior. He smiles or in some other positive way recognizes student contributions.

3. Teacher O.K. (.37) The teacher confirms content; he makes a positive judgment with minimal and terminal acknowledgement.

4. Teacher Confirms. (.81) The teacher incorporates student ideas and/or uses them in lecture. The teacher gives information or an opinion in response to a student verbalization. He is attentive to student

feedback and questions, and may interrupt himself to include student comment, or shift the direction of action to respond to students.

5. Teacher Ponders. (.35) The teacher ponders a student response or expresses doubt. The teacher gives qualified acceptance. He disagrees with a student's response but seeks alternatives. This coding includes asking students if further explanation of previous statements is needed.

6. Teacher Corrects. (.96) A behavior change requested is specified. The teacher corrects or questions what preceded. An opportunity to give the right response or remediation is offered. The teacher may use either a serious or a humorous mode.

7. Teacher Criticizes. (.97) In a minimal manner, student behavior is condemned. A change of behavior is requested, but no second chance is given for the student to make a correction. This includes cold, hostile, sarcastic remarks, and scolding, teasing, and belittling.

8. Teacher Yea. (.80) The teacher praises himself or expresses self-approval. This coding includes the denial of mistakes, and evaluating the correctness of preceding material.

9. Teacher Tangential. (.60) The teacher engages in tangential talk or action to himself. His behavior is not immediately related to the situation. The teacher is preoccupied with something other than teaching.

10. Teacher Owns Up. (.75) The teacher scolds himself, expresses self-disapproval, admits an error, or rechecks his work.

11. Teacher Initiates. (.21) The teacher initiates a probe or asks a broad question, i.e., an open-ended question.

12. Teacher Manages. (.88) The teacher gives procedural directions, or asks narrow (closed) questions with predictable answers. The directions or questions may be either substantive or procedural.

13. Teacher Lectures. (.94) The teacher gives information or an opinion which is not in response and is not feedback. Students, meanwhile, are passive and receptive. This coding includes ignoring student attempts to participate.

14. Teacher Silent Work. (.97) The teacher may be grading papers, writing on the board without reference to students, arranging her material on a bulletin board, or operating a projector.

15. Teacher Delves. (.84) The teacher probes the meaning of a student response. Correctness of student response is not an issue. The teacher asks for more information about his own interpretation, reflection, or incorporation of student idea.

16. Student Zeal. (.48) A student responds eagerly, waves his hand. A student listens attentively. A student accepts,values, or recognizes another's feelings. This coding includes displays of pleasure, appreciation,or good mood, and laughing, crying, or responding emphatically to or with someone.

17. Student Encourages. (0) A student encourages the teacher or another student to go on. This includes thanking another student for help. A student gives approval, praises. Choosing in a game, election, or panel situation is included.

18. Student O.K. (0) A student makes any acknowledgement that the teacher is right (acquiescence) that is not included in another category.

This category includes simple responses such as "Yes, sir."

19. Student Explores. (0) A student asks for information. He may be incorporating a teacher idea in a response. A student gets the teacher or another student to give an idea or talk (task oriented).

20. Student Usual. (.87) A student gives routine feedback in response to a teacher direction, or question, whether the response is correct or not.

21. Student Questions. (.12) A student questions or ponders a preceding response by doubting, arguing, or bringing up new information.

22. Student Suggests. (.92) A student requests a change of behavior and/or makes a correcting suggestion. This may be either serious or humorous behavior.

23. Student Resists. (.85) A student resists. He openly ignores the teacher, e.g., engages in rudeness, hostility, aggressive antipathy, or obvious footdragging.

24. Student Rejoices. (0) A student praises himself or expresses self-approval.

25. Student Woolgathering. (.14) Student extraneous behavior with only the self involved. A student may look bored, yawn, or be sleeping. This coding includes rest periods in primary grades.

26. Student Admits. (.34) A student owns up or admits error. He expresses self-disapproval. This may include actions such as banging a fist on the desk, if the action is clearly self-punitive.

27. Student How. (.65) A student asks for "the" answer. He asks for directions on how to do something without reference to a preceding teacher behavior. The student asks if a preceding answer is right. This

coding includes a student seeking approval or permission to do something.

28. Student Brings Out. (.94) A student gives information or an opinion, or reads a report. A student recites.

29. Student Silent Work. (.97) Activity which is not under immediate supervision (individual or group) such as doing assignments, art work, sharpening pencils, or engaging in computer assisted instruction is coded as Student Silent Work.

Classroom Observation Scales (COS)

The 12 behavioral variables provided by this instrument are listed and described below. Each of these variables corresponds to a particular category of behavior. The score for a variable was the relative frequency of occurrence of the behavioral category in question--i.e., 100 times the frequency of occurrence for the actegory divided by the frequency of occurrence for all categories. The score for a variable is thus interpretable as the percentage of time the behavior in question occurred. The reliability coefficients given for each variable are intraclass correlations based upon two separate sets of data, as reported by Emmer (1971). The first data set involved two observers who each made five observations on each of 15 fifth-grade classrooms. The second data set was based upon observations made by ten observers in 31 first- and second-grade classrooms. In most cases, each classroom was observed four times by two observers.

1. Level of attention. (.62 and .89) Attention as defined by this scale refers to pupil orientation toward the teacher, the task at hand, or whatever classroom activities are appropriate.

2. Teacher-initiated problem solving. (.63 and .73) This variable indicates the degree to which the teacher exhibits a particular style of

instruction. Specifically, teacher-initiated problem solving refers to a pattern of behavior in which the teacher frequently addresses questions and problems to the entire class.

3. Pupil-to-pupil interaction. (.69 and .87) Substantive utterances in which one pupil interacts with another pupil, a group of pupils, or responds indirectly to the teacher are classified on this scale.

4. Teacher presentation. (.83 and .62) By teacher presentation is meant substantive (content oriented) verbal or non-verbal behavior that provides information, and does not imply or require pupil response, nor evaluate pupil behavior.

5. Negative affect. (.67 and .88) This scale includes behaviors that show negative or hostile feelings on the part of either or both the teacher and the pupils.

6. Positive affect. (.64 and .81) This dimension comprises those teacher behaviors that show support of and positive regard for pupils and their behavior.

7. Higher cognitive level student behavior. (.75 and .12) Higher cognitive processes are involved when a student makes a generalization or inference, explains an answer by citing data or rules, solves a problem by combining or using other principles or rules, and defines concepts by citing classes of objects or events, rather than single examples.

8. Passive pupil behavior. (.35 and .69) Withdrawal by the pupil from engagement with his surroundings, visual wandering, and passive observation in which the student avoids maintaining contact for any length of time are considered passive behaviors.

9. Convergent-evaluative interaction. (.82 and .48) Such interaction is characterized by a focus upon obtaining the correct answer to the teacher's question, with little or no attempt to continue the contact once the answer has been obtained.

10. Task orientation. (.13 and .82) This scale is a measure of the degree to which the teacher works toward content-related, substantive goals.

11. Clarity. (.56 and .60) Clarity refers to the degree to which the teacher's presentation of material and his substantive interactions are understood by the pupils.

12. Enthusiasm. (.52 and .68) This variable indicates the extent to which the teacher displays interest, vitality, and involvement in his subject and his instruction.

Adjective Self Description (ASD)

Adjective Self Description (ASD) was administered to subjects as a pre- and posttest criterion measure. The ASD is one of the standard personality and attitude measures used in the COMPASS battery, and represents a concise, direct means of measuring major aspects of self-perception.

The subject is asked to circle one of five numbers on a "No" to "Yes" scale after each of 56 descriptive words to represent how well each word describes the subject. As described in the manual for the ASD (Veldman, 1970), factor analytic procedures have been employed to identify seven basic dimensions of self description. Eight adjectives having the largest factor loadings on each of the seven dimensions were selected to construct the ASD. A number of investigations have been conducted to determine the reliability and validity of the ASD, as reported in Veldman

(1971). Retest stability coefficients for the seven scales obtained using 61 college juniors with a two-week interval ranged from .80 to .92. Alpha coefficients of interval consistency for the items of each scale, as determined on a sample of 713 female junior education subjects, ranged from .64 to .88, the scale with the lowest level of consistency being Ideology.

The scale names and descriptions for the ASD instrument are:

1. Attitude. This scale corresponds to social warmth. A high scale score reflects a positive attitude or high social warmth. Representative items are "cheerful," "gentle," "good-natured."

2. Behavior. This scale corresponds to social abrasiveness or hostility. A high score indicates high hostility. Representative items are "obnoxious," "indifferent," "rude."

3. Efficiency. This scale corresponds to ego organization. A high score indicates high efficiency. Representative items are "efficient," "industrious," "organized."

4. Introversion. A high score on this scale indicates high introversion and low extroversion. Representative items are "quiet," "reserved," "shy."

5. Anxiety. A high score on this scale indicates high anxiety. Representative items are "anxious," "emotional," "moody."

6. Idealism. This scale corresponds to measures of individualism. A high score represents high idealism or individualism. Representative items are "complicated," idealistic," "individualistic."

7. Attractiveness. A high score represents high attractiveness. Representative items are "charming," "good-looking," "sexy."

Career-Related Behavior

Many recent evaluation designs have included unobtrusive measurement techniques, but only a few evaluation reports have actually presented unobtrusively collected data. The ideal unobtrusive measure for evaluating career-related behavior in a teacher education program is a measure as close to on-the-job performance as possible. The period of the present study ruled out the possibility of obtaining ideal data. Within the situational constraints, however, one suitable unobtrusive measure of career-related behavior was devised for use as a posttest criterion.

The strongest behavioral link between a teacher training program and a teaching career is job seeking. The date on which each of the 77 subjects in the present investigation activated a teacher placement file was obtained from records at University of Texas Teacher Placement Service. Neither the subjects nor any staff member working with the subjects were aware that this data was being collected.

The number of days prior to June 1, 1973, the placement file had been activated was assigned as each subject's score. This score, it is proposed, should reflect career motivation.

My Feelings During Videotaping

Five Likert-type items drawn from the State-Trait Anxiety Inventory, State Scale (Spielberger, Gorsuch and Lushene, 1970) were used to assess situational anxiety experienced by student teachers while they were teaching their videotaped lessons. It has been shown (O'Neil, 1972) that this five-item version of the State anxiety scale correlates .84 with the full 20-item scale, which appears to be a valid measure of transitory or state anxiety (Spielberger et al., 1970). Instructions for the scale were adapted

for retrospective and specific reference to the videotape lesson experience.

Immediately after both pre- and posttest videotape lessons, student teachers were asked to indicate how they had felt while they were teaching their lessons. Each of the five item statements (e.g., "I felt calm.") appeared with four response choices varying from "Not At All," to "Very Much So."

Each item response received a score ranging from 1 to 4. Where items were negatively phrased, scoring was reversed, so that higher scores always indicated lower levels of anxiety. Item scores were then summed to yield a single score for each student teacher.

Our Lesson

Our Lesson was constructed by the evaluation staff for administration to pupils at the conclusion of the videotaped lesson session. Six items were constructed to sample pupil evaluation of lesson content and presentation. These items were designed to tap variables frequently cited in the literature on teacher effectiveness as correlates of pupil learning: pupil interest, lesson clarity, appropriateness of instructional level, teacher enthusiasm, pupil enjoyment, and opportunity for pupil response. Two other items pertained to the representativeness of the lesson situation, and were included to elicit pupils' opinions as to whether their behavior as well as their student teacher's behavior was "about the same" as in the regular classroom.

Pupils were required to respond to each of the eight item statements by making an X over a smiling face if they agreed with the statement, over a frowning face if they disagreed, and over a neutral face if they had no opinion. All items were positively phrased.

Our Lesson was administered during pretest data collection only. The scores for each student teacher used in the analysis of data were mean pupil responses for each item on a 1 to 3 scale (disagree = 1, neutral = 2, agree = 3).

Professional Plans and Affiliations Questionnaire

The Professional Plans and Affiliations Questionnaire is a self-report instrument which was constructed by the evaluation staff for use as a posttest instrument only. Many of the items for the instrument were taken from the Modified Exit Interview Questionnaire, which was administered as a pretest instrument, and readministered in either exact or modified form.

The first two items on the questionnaire were queries as to the subjects' intended graduation date. The remaining items appeared in three sections: Ratings, Additional Comments on Teacher Training, and Teaching Motivation and Plans.

Ratings. Subjects were asked to rate their public school supervising teacher, in comparison with all public school teachers they had ever encountered, on the following dimensions: (a) Understanding, friendly, (b) Responsible, businesslike, (c) Stimulating, enthusiastic, (d) Helpful to me, (e) Interested in me, and (f) Concerned with children. Response options for each characteristic were: "Much More," "Somewhat More," "About the Same," "Somewhat Less," or "Much Less." Responses were scored on a 1 ("Much Less") to 5 ("Much More") scale.

Similarly, subjects were asked to rate their student teaching semester college supervisors in comparison with all other college instructors they

had encountered. Response options and scoring were the same as for the public school supervising teacher rating, but the rated characteristics differed slightly. They were (a) Understanding, friendly, (b) Responsible, businesslike, (c) Stimulating, enthusiastic, (d) Helpful to me as a teacher, (e) Helpful to me as a person, (f) Interested in me, and (g) Provided specific suggestions.

Next, subjects were asked to rate six courses in the professional sequence, including student teaching, as "Highly Essential," "Somewhat Essential," "Average," "Somewhat Irrelevant," or "Very Irrelevant." Responses were once again scored on a 1 ("Very irrelevant") to 5 ("Highly Essential") scale.

Two open-ended questions completed the Ratings section of the questionnaire. Subjects were asked what they considered to be the most valuable developmental college experience they had had in terms of preparing them to teach, and were asked what, in their entire teacher preparation, they regarded as having been the greatest waste of time. Responses were never coded due to time constraints.

Additional Comments on Teacher Training. This section of the questionnaire consisted of one "Yes"-"No" and four open-ended questions. Student teachers were asked to name one thing they would like to see receive increased emphasis in their teacher preparation program, one thing they would like to see changed or dropped, and one thing they would like to see added to their teacher education program. Finally, they were asked to check "Yes" or "No" to the question, "Do you feel that you have benefitted or gained anything from your student teaching experience?" If they responded positively, they were asked to elaborate on the response by indicating specific ways they had benefitted. None of the responses in this section of the

questionnaire were coded due to time constraints.

Teaching Motivation and Plans. A series of four questions was designed to assess subjects' current teaching motivation, and their impressions of how their motivation to teach may have changed during the student teaching semester.

Subjects were asked to rate their current motivation to teach on a 1 ("Zero") to 5 ("Very Great") scale. Next, they were asked to indicate whether, as teachers, they thought they would be "Exceptionally Good," "Above Average," "Average," "Fair," or "Poor." This response was also scored on a 1 (low) to 5 (high) scale. Student teachers were asked whether, during the student teaching semester, they felt their motivation to teach had "Increased a Great Deal" (scored 5), "Increased Somewhat" (4), "Remained About the Same" (3), "Decreased Somewhat" (2), or "Decreased a Great Deal" (scored 1). An open-ended question followed immediately, asking, "If your motivation to teach has changed during the student teaching semester, to what do you attribute the change?" Responses to this question were coded into three categories, attributing the change to (1) knowledge and experience gained in student teaching, (2) positive personal experience with the public school or college supervisors, or other instructors, or (3) a negative experience in student teaching (either situational or interpersonal).

Five questions assessed plans for the future. Subjects were asked whether they ever planned to take any job outside the home other than teaching. Response options, scored on a 1 to 5 scale, were: "Yes," "Probably," "Undecided," "Probably Not," and "No." Next, subjects were asked to check the number of years they thought they would teach. The response categories,

also coded on a 1 to 5 scale, were "0," "1," "2 to 5," "6 to 10," and "More than 10." Another uncoded open-ended question followed, asking those who never planned to teach what sort of job they would like to have.

In the next question subjects were asked how soon they planned to teach. Response options ranged from "As soon as I graduate" (scored 5) to "Never" (scored 1). Subjects were also asked whether they planned to teach if they were married. Response options were "Yes," "Probably," "Undecided," "Probably Not," and "No," scored 5 to 1 in order.

The tenth item in this section called for listing the professional organizations or societies to which subjects belonged. Item 11 asked for a list of professional magazines or newsletters to which subjects subscribed. These questions were scored by simple numerical counts.

The last two items on the questionnaire probed current job status. Subjects were asked whether they had applied for a teaching position yet, and if they had, were asked to indicate how many applications they had made and how many interviews they had had. Finally, subjects were asked to check their current job status as, "I have been offered or have accepted a teaching position" (scored 3), "I am currently seeking a teaching position" (scored 2), or "I am not seeking a teaching position at this time" (scored 1).

Examination of response distributions for items in the Teaching Motivation and Plans section of the questionnaire eliminated some items from further consideration. As indicated above, response opportunities varied across the items. The original scoring of some of the items was altered after the distribution on each item was examined. In all cases this modification consisted of reassigning score units after adjacent response opportunities of low frequency were combined. Factor analysis of the remaining items

resulted in two factors: Factor I, Job Seeking Behavior, and Factor II, Motivation to Teach. The items which loaded on Factor I were those asking the subject the extent to which she had applied, been interviewed, or obtained a job. Items loading on Factor II were those asking the subject the extent of her commitment to a teaching career. Factor scores were calculated for each subject for these two factors and were used in subsequent analyses. Table 2-1 presents the item stems and their factor loadings.

Table 2-1

Varimax Factor Loadings, Communalities, and Percentages of Variance Accounted for by Two Factors Derived from the Professional Plans and Affiliations Questionnaire

Item Stem	Varimax Endings		h^2
	Factor I	Factor II	
Have you applied for a teaching position yet?	-.92	-.11	.86
Current job status	.83	.14	.71
Present motivation to teach.	.36	.60	.50
How many years do you plan to teach?	.04	.82	.68
Number of subscriptions to professional literature.	.11	.68	.48
Number of teaching position applications	.87	.23	.81
Number of teaching position interviews.	.83	.16	.72
Percentage of Variance	44.73	23.02	

Note.--Response opportunities varied across the items. The score of a subject on an item was represented by a number assigned after the distribution on each item was obtained, examined, and modified by combining adjacent response opportunities of low frequency.

Profile of Learning Priorities, Form C (PLP)

The Profile of Learning Priorities (Peck, 1972b) was constructed for use as both a pretest and a posttest measure. Items on the instrument were directed at the three domains of competence outlined in the Introduction to this report: Personal, Interpersonal, and Instructional. Due to time constraints, only data from the Instructional Domain portion of the PLP was analyzed for inclusion in this report.

To complete the PLP, the subject was asked to read each of a series of descriptions and rate himself on each description on a 1 to 5 scale. The subject was instructed to circle the 5 if he considered himself to be in the top 20% of all student teachers he had known, 4 if he considered himself to be among the next highest 20%, and so on.

Posttest responses of 77 subjects to the 16 Instructional Domain items were factor analyzed. Four factors were extracted using a principal components analysis and a Varimax rotation procedure. The four factor names, along with representative items for each are given below.

I. Competent Management. "Maintains adequate classroom discipline; minimum disturbances." "Direct, immediate action to solve teaching problems; no procrastination or avoidance." "Successfully identifies workable solutions to teaching problems (independently or group)." "Maintains orderly group movement toward goals the children understand." The percent of variance accounted for by this factor was 22.75.

II. Flexibility. "Flexible in adapting plans to new circumstances." "Flexible in conducting class; adjusts well to new or unusual situations." "Diagnoses learning needs of individual children perceptively and accurately." "Resourceful in drawing on people and materials in planning instruction." This factor accounted for 18.43% of the total variance.

III. Professionalism. "Feels a personal interest in being a teacher; likes it, respects it, and intends to pursue it as a career." "Physical facilities, instructional aids, and time are well-organized." "Realistically aware of own teaching behavior and attitudes." This factor accounted for 13.67% of the variance.

IV. Responsibility. "Feels personally responsible to maximize children's learning; does not blame the children, home, or society for all failures to learn." "Achieves an accurate, thorough knowledge of the subjects to be learned." This factor accounted for 10.10% of the total variance.

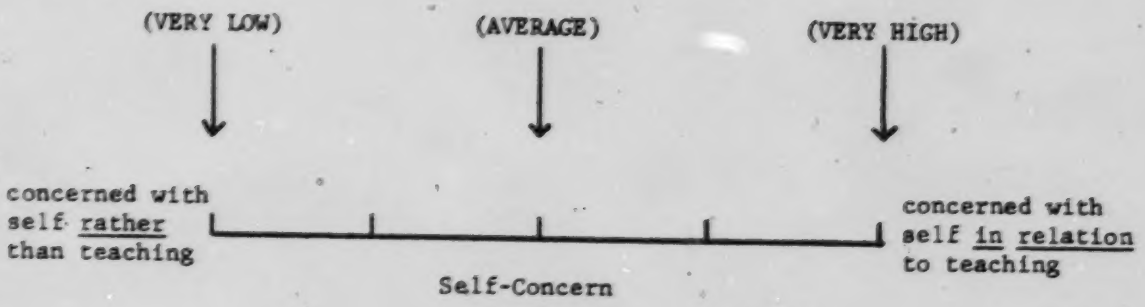
Readiness Assessment

This instrument was constructed for the purpose of obtaining ratings of "readiness to teach" for each student teacher participating in the study. Readiness Assessment was used for self-ratings by the student teachers and for other-ratings by the public school supervising teachers, college supervisors, and PTE counselors in both pre- and posttest data collection.

"Readiness to teach" was conceived as a summary variable having a number of contributory dimensions. Readiness Assessment contained items representing six such dimensions: self-concern, concern for children, concern for impact, perceptive about self, perceptive about children's behavior, and motivated to teach.

Each item appeared as a scale continuum, and respondents were asked to mark the point on the scale which corresponded to their rating of a particular student teacher in relation to all student teachers they had known. The continuum scale ranged from "very low" through "average" to "very high," as

shown below:



As indicated in the example, the extremes on the continuum were given along with the dimension name for further response guidance.

Responses were assigned scores ranging from 1 (very low on the continuum) to 4 (very high on the continuum). Analyses were performed separately for each item score, for the posttest administrations only.

Self-Report Inventory (SRI)

Like the ASD, the Self-Report Inventory (SRI) is a COMPASS battery instrument administered as a pre- and posttest criterion variable.

The SRI contains 48 items, each of which consists of a statement followed by a five-choice scale ranging from "like me" to "unlike me." Six items are assigned to each of eight scales. Thus, the instrument yields eight scale scores, which have a lower limit of 6 and an upper limit of 30. The scale names and descriptions for the SRI are:

1. Self. A high score represents a positive attitude toward self. Example item: "In almost every respect, I'm glad to be the person I am."

2. Others. A high score represents a positive attitude toward others. Example item: "The way I get along with my friends is extremely important to me."

3. Children. A high score represents a positive attitude toward children. Example item: "I'm very comfortable and happy when I am with children."

4. Authority. A high score represents a positive attitude toward persons in authority. Example item: "I really enjoy getting to know people in positions of authority."

5. Work. A high score indicates a positive attitude toward work. Example item: "Doing a good job in anything I undertake is very important to my sense of well-being."

6. Reality. A high score represents a positive attitude toward reality. Example item: "I live in accordance with the idea that 'It is better to have loved and lost than never to have loved at all.'"

7. Parents. A high score represents a positive attitude toward parents. Example item: "I am very happy with my present relationship with my parents."

8. Hope. A high score represents a positive attitude about the future. Example item: "I really look forward to the time when I will be settled down to my life's work."

Detailed information on the SRI is available in the manual (Bown and Veldman, 1967). Alpha coefficients of internal consistency obtained using a sample of 2321 freshman college students as reported in the SRI manual, are: Self, α = .78; Others, α = .65; Children, α = .85, Authority, α = .53; Work, α = .70; Reality, α = .28; Parents, α = .84; and Hope, α = .66.

Student Evaluation of Teacher 2

A 23-item version of the Student Evaluation of Teacher 2 (SET 2) instrument was administered to the pupils of student teachers near the end of the spring semester 1973, as a posttest criterion measure. A manual for the SET 2 has previously been published (Haak, Kleiber, and Peck, 1972). Briefly, the SET 2 asks the pupil to respond in a "true" or "false" fashion to 23 items describing his teacher. Example items are: "She makes school fun," "She helps us a lot," "She gets mad a lot," and "She likes me." For each student teacher, pupils' responses to each item were averaged, thus reducing the data for each student teacher to 23 item mean scores.

Previous versions of the SET 2 have been carefully studied with respect to factor structure (Haak et al., 1972). Since the version of SET 2 used in the present study differed slightly from previous versions, in that it incorporated one new item and one revised item, it was necessary to factor analyze the data obtained. A Varimax Rotation procedure (Veldman, 1967) was employed to extract five factors based on the 23 items. The three factors accounting for the largest percentage of variance (47.7%) were chosen. These three factors correspond closely to those derived by Haak et al. (1972). Factor names and representative items are:

Rapport. The lower the score on this factor, the higher the rapport between student teacher and pupils. Example items for this factor are: "She likes us kids;" "She makes school fun;" and "The kids like her."

Unreasonable Negativity. The lower the score on this factor, the greater the unreasonable negativity of the student teacher. Example items are: "She thinks I act ugly;" "She thinks I am lazy;" and "She gets mad a lot."

Fosterance of Self-Esteem. The lower the score on the factor, the greater the student teacher's fosterance of pupil self-esteem. Example items are: "She likes for me to help her;" "She thinks I am smart;" and "She thinks I work hard."

Student Evaluation of Teacher Training Program

Ten objectives were selected from the Basic Program Plan for the PTE Program and used as constructs for the instrument named Student Evaluation of Teacher Training Program. A pool of Likert-type items was generated for each construct. From this pool, four to seven items were finally selected for each construct, resulting in a collection of 55 items comprising ten scales.

The instrument was administered to student teachers during both pre- and posttest data collection. Alpha coefficients of internal consistency were computed for each scale based on the sample of 71 student teachers pretested with the instrument. The range of alpha-reliabilities was from .68 to .84, with the average reliability being .78.

The ten constructs measured are:

Scale 1: Program Integration. The program is an integrated system, not merely a collection of course units. Knowledge gained in one course transfers to others. Concrete experiences relate to theory throughout the total program.

Scale 2: Individualized Teaching. The instructors adapt their teaching to individual needs. The program is itself a model for individualization of instruction.

Scale 3: Aid to Autonomy. The program stresses the importance of realistic self-confidence and professional competence as bases for decision-

making in teaching situations.

Scale 4: Constructive Feedback. The program provides continual intensive feedback to aid prospective teachers in developing effective behaviors and modifying inappropriate behaviors.

Scale 5: Behavior Modeling. The instructors in the program identify strongly with teacher training as a personally and professionally rewarding role. Through their behavior they impart a value for teaching and motivate students toward professional goals.

Scale 6: Negotiation. Within the constraints of established College of Education requirements, the program provides opportunities for negotiation of learning methods and curriculum flexibility. The student experiences involvement and participation in planning her own learning experiences.

Scale 7: Teacher Preparation. The program provides the content knowledge and learning experiences needed to develop necessary teaching behaviors. The students feel prepared to teach.

Scale 8: Teacher Educator--Student Teacher Interaction. Personal interaction between student teachers and their teacher educators is an important component of the program. Teacher educators invite, initiate, and sustain personal dialogue with their students.

Scale 9: Person-Centered. The prospective teacher is the focus of the program. Her concerns, development, and goals are the core of the program.

Scale 10: Personal, Intellectual, and Social Development. The entire program provides an opportunity and the support for prospective teachers to develop personally, socially, and intellectually.

Teacher Beliefs

Wehling and Charters' work in the area of teacher beliefs (Wehling and Charters, 1969) resulted in a carefully developed, research-based instrument for measuring teacher beliefs. During the pretest phase of data collection only, subjects in the present investigation completed a subset of items from the Wehling and Charters instrument. This 46-item subset contained all items on three scales which were selected as operationalizing constructs which the PTE Program seeks to effect. These three scales are defined as follows:

Scale 1: Personal Adjustment Ideology. The belief that the instructional process should be organized around student needs and interests in order to contribute to social and emotional development. This is a cognitive belief in an idea, not an inclination to establish warm relationships with pupils.

Scale 2: Student Autonomy versus Teacher Control. The belief that the locus of control in the classroom should lie with the students. This belief expresses the amount of faith a teacher has in students and their capacity for spontaneous learning.

Scale 3: Consideration of Student Viewpoint. The belief that empathy is an instructional strategy and the teacher must be sensitive to pupil feelings and display friendliness. This has affective as well as cognitive components in contrast to Scale 1.

Six response alternatives were available for each item statement, ranging from "Strongly Agree" to "Strongly Disagree," and were scored from 1 to 6. Scale scores were derived by unit-weighting those items which loaded on each scale (Wehling and Charters, 1969). Only the Student Autonomy Scale involved negative loadings and thus, reverse weighting.

Pretest data analyzed for Teacher Beliefs consisted of three scale scores for each student teacher.

Teacher Concerns Checklist.

The Teacher Concerns Checklist (TCCL) was administered to subjects as a posttest criterion measure. The purpose of the TCCL is to determine about what things teachers are most concerned, this instrument having been developed in the context of Fuller's Concerns Theory (Fuller, 1969a). The TCCL consists of 56 Likert-scaled items, each item presenting the statement of a possible concern (e.g., "I am concerned about lack of respect of some students."). For each item, the subject is asked to check one of five blanks corresponding to the extent to which he or she possesses the concern presented in the item. The five response alternatives are "Not concerned at all," "Slightly concerned," "Moderately concerned," "Very concerned," and "Extremely concerned"--these response alternatives being coded as 1 through 5, respectively.

Factor analyses of the TCCL (Watkins, 1973) has yielded three basic subscales, and these three scales correspond to the three basic levels or stages posited in the Concerns Theory. In the same report, Watkins presents results relevant to the reliability of the TCCL subscales. Test-retest reliability coefficients ranged from .69 to .75 for one sample (58 undergraduate education majors) and from .80 to .83 for a second sample (44 subjects who were either undergraduate education majors or student teachers). Alpha coefficients of internal consistency for the items of each scale ranged from .71 to .89 for one sample (95 undergraduate education majors), .81 to .94 for a second sample (262 undergraduate education students and 73 student teachers), and .82 to .93 for a third sample (345 inservice teachers).

Descriptions of the three scales and representative items are presented below:

1. Self-Concern. This scale represents concerns about comfort, adequacy, or success as a teacher--a high score indicating greater concern. Items loading most heavily on this factor are: a) "I am concerned about feeling more adequate as a teacher," b) "I am concerned about whether the students really like me or not," and c) "I am concerned about how students feel about me."

2. Professional Concerns. This scale represents concerns about the demands upon teachers, teaching circumstances, and professionalism--a high score indicating greater concerns. Items loading most heavily on this factor are: a) "I am concerned about (the fact that) the mandated curriculum is not appropriate for all students," b) "I am concerned about (being) frustrated by the routine and inflexibility of the situation," and c) "I am concerned about feeling under pressure too much of the time."

3. Concern for Pupils. This scale represents concerns for pupils as learners--a high score indicating greater concerns. Items loading most heavily on this factor are: a) "I am concerned about (the) slow progress of certain students," b) "I am concerned about adapting myself to the needs of different students," and c) "I am concerned about helping students to value learning."

Trait Variables

The trait variables in the present investigation were measured using three instruments from the Comprehensive Personal Assessment (COMPASS) Battery, which had been administered to subjects during their first semester of teacher training.

Specifically, all scales of the Adjective Self-Description, Self-Report Inventory, and One Word Sentence Completion instruments were used

to measure trait variables. The Adjective Self-Description and Self-Report Inventory were also administered in the spring semester of 1973, as criterion variable sources, and are described above in the section on criterion variables.

One-Word Sentence Completion.

One-Word Sentence Completion (OWSC) is a 62-item projective instrument developed at the R & D Center in connection with the Computer Analysis of Personality project, supported by NIMH. Subjects are asked to respond to sentence stems by completing the sentence with a single word. Responses are handwritten in blanks on the protocol. Example stems are: "I enjoy _____ very much," "I am afraid of _____," and "Children usually _____ me."

A number of computer-based scoring systems have been constructed for the OWSC and used in different research projects (Veldman, Menaker, and Peck, 1969; Veldman and Bown, 1969; Veldman, 1970; Veldman, 1973). For the present investigation, eight of the scales from Veldman's most recent scoring system (Veldman, 1973) were employed. Responses were prepared for scoring in the following manner: 1) misspellings were corrected, 2) punctuation, spaces between multiple words, and initial articles were removed, 3) the length of a response was limited to 10 characters (for responses involving more than ten characters only the first 10 characters were retained), 4) proper names were coded "PN."

The eight OWSC scales employed in the present investigation were as follows:

1. Response Length. This variable is the mean number of characters per response excluding blanks, proper name codes, and numeral responses.

2. Repetitions. This variable indicates the number of repetition responses in the protocol. The response to an item is counted as a repetition response if the same response occurs for one or more other items. Note that if the response "good" is given to three items, then three repetition responses are involved. Blanks, proper name codes, and numerical responses were never counted as repetition responses.

3. Popular. This variable is the number of responses that are classified as popular responses. A response is defined as popular for a given item if more than 10% of the normative sample gave that response for the item. The normative sample was that reported by Veldman (1971) and included 1718 students enrolled in the introductory educational psychology course (junior level) at the University of Texas at Austin during the fall 1968, spring 1969, and summer 1969 semesters. This sample included 341 males and 1377 females, with the female students including 457 elementary and 920 secondary majors.

4. Evasion. This variable is the number of evasive responses in the protocol. Evasive responses include failures to respond to the item (blanks), the use of private proper names, repetition of a key word in the item (e.g., "Darkness is dark."), cryptic references (e.g., "I hate him."), apparently deliberate ambiguity (e.g., "Most men are different."), and numeral responses (usually ages).

5. Hostility. This variable is the number of hostile responses in the protocol. Hostile responses are those that suggest antagonism toward or devaluation of other people, projection of blame for personal dissatisfaction, or authoritarian attitudes.

6. Anxiety. This variable is the number of anxious responses in the protocol. Anxious responses are those that indicate apprehension, self-doubt, unusual fears or abnormal tension. On some items, an apparent projection of anxious feelings to others is also defined as an anxious response (e.g., "Most men are anxious.").

7. Depression. This variable is the number of depression responses occurring in the protocol. Depression responses indicate self-derogation, loneliness, or depression. On some items, an apparent projection of feelings of depression to others is also counted as a depression response.

8. Rejection of a Teaching Career. This variable indicates the number of responses in the protocol which indicate rejection of teaching as a career.

Additional Instruments

A number of instruments were administered but not analyzed and/or not reported here, either because of the lack of time or because the distributions for these instruments were truncated or highly skewed. Since no results are reported for these instruments, they are described only briefly, but with the intent that others may find them useful.

Career Information Form.

This form, administered as a pretest instrument, was used to obtain biographical information from the subjects. Additional sections of the questionnaire asked subjects for details of any previous teaching experience, a brief work history, a description of personal interests (i.e., hobbies, organizations, or other activities),and a short self-assessment. The self-assessment section consisted of three open-ended questions asking the subject to describe his greatest personal strengths and limitations as they might relate to teaching effectiveness, and give any other comments about himself that he wished.

Directed Imagination.

Directed Imagination is a timed projective instrument which was administered as a pretest only. The subject is asked to write a series of four fictional stories about teachers and their experiences. The subject is given four minutes to write each story.

Group Atmosphere Rating.

The Group Atmosphere Rating was administered as a pretest only, for the purpose of assessing subjects' feelings about both the ideal group atmosphere for a teaching training program, and the actual group atmosphere in the training program they had been going through. A series of twelve bipolar adjectives (e.g., tense-relaxed, warm-cold, closed-open) appeared at the extremes of a seven-point scale. Subjects were asked to indicate the ideal and actual group atmospheres by placing an X and a ✓, respectively, at the appropriate points on each of the twelve scales.

Individual Locator Form.

This form was completed at pretest and updated at posttest time. The subject was asked to give his campus or other Austin mailing address, and to supply the name and address of a friend or relative for the purpose of mailing the subject a summary of the results of the present study when available.

Interpersonal Skills Questionnaire, Form P.

This instrument was administered as a posttest measure of subjects' attitudes towards their behavior with others. The subject was asked to rate each of 20 statements on a five-point scale from "Not at all true of myself" to "Completely true of myself." Example statements are "I am

confident of myself," "I say what I feel," "I lead groups effectively," and "I know who I am."

Modified Exit Interview Questionnaire.

This questionnaire was administered as a pretest measure only, although many of the items were readministered in the same or modified form as part of the posttest instrument Professional Plans and Affiliations Questionnaire, which is described in detail above as a criterion variable source. The subject was asked his intended graduation date and details of his degree program. Ratings of college and public school supervisors during the observation semester, and ratings of selected teacher training courses were obtained. Three open-ended questions queried the subject as to the most and least valuable college experiences in terms of teacher preparation, and most valuable college experience in terms of personal growth. A final section of the questionnaire assessed teaching motivation and plans.

My Teacher Education Instructors.

During pretest data collection subjects were asked to respond to this series of seven statements by indicating whether each statement was more or less characteristic of the teacher education faculty they had experienced. Responses ranged on a five-point scale from "Definitely Untrue" to "Definitely True." Example item statements are "Most of my teacher education instructors were interested in me as an individual," and "In our teacher education courses students were encouraged to think for themselves."

Our Lesson (Observer Rating).

This form replaced the original form of Our Lesson, and was independently completed by the three members of the videotape crew for each

student teacher who made a posttest videotape. Each rater appraised the lesson as being new or review material and responded to seven statements on a five-point scale from "Definitely No" to "Definitely Yes." Four of the statements corresponded to items used with pupils on the original Our Lesson instrument. The three new item-statements were "The information given by the teacher was correct," "The teacher knew her subject content," and "The teacher effectively presented the subject content to the pupils."

Profile of Learning Priorities, Form A.

Form A of the Profile of Learning Priorities was completed by public school supervising teachers as a pre- and posttest form. Instructions asked for an other-rating of each student teacher as on Form B for counselors and college supervisors. There were fewer items on Form A, however, since time constraints made it impossible to administer all items appearing on the other two forms to the public school teachers. Thus, Form A included 9 items for the Personal Domain, 9 items for the Interpersonal Domain, and 10 items for the Instructional Domain.

Profile of Learning Priorities, Form B.

This instrument is identical in form to the Profile of Learning Priorities, Form C, which is described above as a criterion variable. Only the directions for Form A differed, in that they were written specifically for use by the college supervisors and counselors, who rated each of their student teachers on this instrument at both pre- and posttest time.

Student Teacher Rating.

Three scales from Adjective Self-Description (ASD) were completed by counselors and college supervisors as pre- and posttest other-ratings

of the subjects. All adjectives for the Anxiety, Efficiency, and Attitude scales were used as other-rating stimuli, and appeared in the same format as the full ASD.

Student Teacher's Evaluation of Videotape Session.

This instrument consisted of four questions designed to help assess the subjects' feelings about the respresentativeness of the pretest videotaped lessons. The subject was asked (1) whether, in her opinion, the lesson he had just taught was representative of her usual teaching; (2) to describe specific ways in which she felt the lesson was unrepresentative, if she felt the lesson was not representative; (3) to indicate how much preparation she had done as compared with her normal preparation for regular classroom lessons; and (4) to indicate whether or not anything about the way the videotaped lesson assignment was conducted, or any condition during the taping led to her teaching in a way unnatural for her, and if so, to specify the conditions she felt caused the unnatural teaching.

For the posttest videotaping administration, the questionnaire was revised slightly. In addition to comparing preparation time with regular lesson preparation time, the subject was asked to indicate the approximate amount of time spent preparing the videotaped as lesson. Response choices were: "Less than 15 minutes," "Around 15 minutes," "About half an hour," "Close to 45 minutes," and "More than 45 minutes." A final question, asking the subject how many times prior to the spring semester 1973 she had been videotaped while teaching a lesson, was also added for posttest administration.

Chapter 4

DATA COLLECTION PROCEDURES

Pretest and posttest data for the study were collected from four major sources: a) instruments completed by the student teachers, b) instruments completed by counselors and supervisors, c) videotapes of lessons taught by trainees, and d) data collected from University records. Procedures will be described under these four headings.

Information from UT Records

A requirement of the University of Texas College of Education specifies that all students entering a teacher training sequence must complete the Comprehensive Personal Assessment Battery designed to measure the attitude and personality characteristics of prospective trainees. The Battery is normally administered to teacher trainees in the first two weeks of the junior year, during a regularly scheduled class period of introductory educational psychology. For this study baseline scores on instruments pertaining to the assessment battery, the dates on which subjects opened placement files at the University Teacher Placement Service, cumulative grade point average after six semesters of college work, and area of academic specialization were obtained from College of Education records for each subject.

Instruments Completed by Student Teachers

Pretest. Pretest data collection for student teachers at all schools other than School 1 was completed within the first month of the spring semester, 1973. At schools 2, 3, and 4, instruments were completed in two

sessions conducted by one of three evaluators. During the first session, scheduled for a two-hour block of time, subjects filled out Computer ID Cards, Participation Consent Form, Individual Locator Form, Directed Imagination, Adjective Self Description, Self-Report Inventory, One Word Sentence Completion, Career Information Form, and Modified Exit Interview Questionnaire.

A standardized set of instructions was read to subjects by the evaluator in charge. These instructions included a description of the purpose of the study and an explanation of the procedures which would be employed to insure confidentiality. Each subject was given a computer card preprinted with a unique student identification number to be used during the study in order to preserve confidentiality. Subjects were asked to print their names on one portion of the card, alongside the number, and return it to the evaluator. They were asked to tear off and retain a second portion of the card, also preprinted with the unique ID number, for their own reference. Completed instruments were processed by removal of the name and substitution of the unique ID number obtained from a central file.

After subjects had been assigned ID numbers, they were asked to fill out the Individual Locator Form, and to label all other instruments with their ID numbers. The Directed Imagination instrument was completed first using the standard four-minutes-per-story timing. The remainder of the instruments for the session were then briefly described, and suggested approximate completion times were given. Subjects were asked to read and use the standard written directions on each instrument, and to pace themselves. Subjects were further instructed to raise their hands if they had questions, and to refrain from talking among themselves during the session.

During the second session, scheduled for a one and one-half hour time period, subjects completed the Readiness Assessment, Profile of Learning Priorities--Form C, Teacher Beliefs, Student Evaluation of Teacher Training Program, Group Atmosphere Rating, and My Teacher Education Instructors. As before, standardized instructions were read to the subjects. General instructions for the session remained the same, and most of the instruments were self-paced, as in the first session. The Readiness Assessment was filled out in group form, with revised instructions calling for self-rating read to the group. Similarly, the Profile of Learning Priorities was filled out as a group, since it was felt that the instructions were complex and required verbal elaboration. The remaining instruments were generally described, and subjects were asked to devote the rest of the session to completing the instruments at their own pace, according to the standard written instructions appearing on each.

Some subjects at School 4 missed the initial test session and were given individual make-up sessions at the R & D Center at their convenience during the following two weeks. Similarly, several School 4 student teachers failed to appear for the second instrument session. When these subjects indicated their unwillingness to schedule make-up sessions, they were allowed to take instrument packets home, fill out the forms as they had time, and return the packets no later than one month into the semester. A written version of the standardized instructions read to all other subjects was included in these packets.

Data collection at School 1 did not proceed as smoothly. A number of subjects did not attend the first scheduled data collection session. Some of the subjects who did attend were reluctant to proceed with the

session as planned. Most of the subjects present completed Computer ID Cards, Participation Consent Forms, Adjective Self Description, or some part of the above. Since it was apparent to the evaluator that additional liaison work was needed, it was agreed 1) that subjects would take home the remainder of the first-session instruments, fill them out as time permitted, and return them to the second scheduled testing session, and 2) that during the second scheduled session, no testing would take place, but a representative of the PTE program would be present to hear and attempt to remedy any objections to participating in the study. Before adjourning the session, the evaluator read those portions of the standardized instructions which were applicable.

At the time originally scheduled for the second data collection session, an administrator for the PTE program met with School 1 student teachers as agreed. It was apparent that by the time further group data collection sessions could be scheduled, the data would not be appropriate for making baseline comparisons. For this reason, take-home packets were brought to the session for distribution. The program administrator heard complaints, most of which seemed to center around the pressures the student teachers were feeling due to time and heavy work demands, and explained the rationale and importance of the study to the subjects in greater detail than was given in the standard instructions. At the end of the meeting individualized instrument packets containing all pretest instruments were handed out. For subjects who had attended the first session, duplicates of the first-session instruments were enclosed along with second-session instruments. This seemed advisable since the evaluation unit had received several reports of groups

of subjects filling out instruments together. Subjects who had not attended the first session were given packets containing all materials for both the first and second scheduled sessions. Written instructions, kept as close as possible to the instructions used orally at other sites, were enclosed in each packet. It was necessary to ask subjects to attempt to time themselves on Directed Imagination, the only timed instrument in the battery.

Posttest. Posttest data collection was carried out in two-hour sessions beginning the last week in April, approximately three weeks before the end of the spring semester. Interpersonal Skills Questionnaire--Form P, Readiness Assessment, Professional Plans and Affiliations Questionnaire, Student Evaluation of Teacher Training Program, Adjective Self Description, Profile of Learning Priorities, Self Report Inventory, and Teacher Concerns were completed as posttest instruments.

Individualized instrument packets were prepared which included 1) a Xerox copy of the Individual Locator Form filled out at pretest time, 2) a copy of the Participation Consent Form only for subjects who had never returned a signed form, and 3) the posttest instruments, packaged in the order they are listed above, and prelabeled with ID numbers.

Subjects were instructed to update the Individual Locator Form if necessary, and were asked to sign the consent form if one had been inserted in their packets. Subjects were reminded they had filled out most of the instruments before, but were asked to read all the standard instructions appearing on the instruments. Instructions for the Profile of Learning Priorities were also summarized in the standard instructions. As in previous instrument sessions, subjects were asked to pace themselves through the instruments, to refrain from talking, and to raise their hands when they had questions.

At School 2 only, instructions were read to the subjects. At the other three sites, due to scheduling problems, it was deemed advisable to include a written version of the instructions in each subject's packet. Thus, as subjects at these three schools arrived for the session, each subject was handed his individualized packet, asked to read the enclosed instructions before starting on the instruments, and directed to a seat to begin working on the instruments.

All School 2 subjects attended the posttest session. The few subjects who missed the sessions at Schools 3 and 4 had indicated they no longer wished to participate in the study. At School 1, seven subjects did not appear for the session. A make-up session was held on May 3rd, at which time several more subjects completed posttest instruments. The remaining School 1 subjects were allowed to take their packets home for completion. The PTE counselors at School 1 distributed and collected these take-home packets, which were completed in several cases as late as May 14th, in the final week of the semester.

SET II. The Student Evaluation of Teacher instrument (SET II) was administered to the pupils of student teachers during the first and second weeks after the student teachers had left the schools. The instrument was administered according to the procedures outlined in the manual for the SET II (Haak, Kleiber, and Peck, 1972).

Instrument Completion by Counselors and Supervisors

Pretest. Pretest collection of data from UT (college) supervisors, public school supervising teachers, and R & D Center counselors took place during the fourth week of the spring semester. At that time the student teachers had been working with their public school supervising teachers in their assigned schools for three weeks.

Rating forms for the counselors and supervisors were distributed with cover letters requesting that the forms be returned to assigned collection agents by the end of the fifth week of the semester. Standard written instructions on each instrument were used. These personnel were asked to call one of the evaluators, whose name and phone number was given, in case of questions.

Program counselors rated PTE student teachers on the Readiness Assessment, Profile of Learning Priorities--Form B, and the Student Teacher Rating.

University of Texas supervisors rated each of their student teachers using the Readiness Assessment, Profile of Learining Priorities--Form B, and the Student Teacher Rating.

Public school supervising teachers filled out the Readiness Assessment and the Profile of Learning Priorities--Form A for the student teachers assigned to their supervision.

There were no major difficulties in collecting the data. However, several of the public school supervising teachers at Schools 1 and 4 cooperated minimally, and a substantial number failed to fill out the Profile of Learning Priorities correctly. Directions for the instrument were therefore revised for greater clarity prior to posttest data collection.

Posttest. The same procedures and instruments were employed for posttest data collection, with the exception of the slightly revised instructions which appeared on the Profile of Learning Priorities--Form A for the public school supervising teachers. The data collection took place over a three-week period ending one week before the semester's end.

Rating forms for the public school supervising teachers were distributed April 30th, and collected May 3rd, for teachers at Schools 3 and 4.

For Schools 1 and 2, these forms were distributed April 20th, and returned on April 26th.

Forms were distributed to UT supervisors working with School 3 and 4 student teachers and to program counselors working with the same student teachers on April 30th, and were returned by May 11th. Forms for the remaining counselors and college supervisors were distributed April 20th, and returned by May 4th.

As expected in such a large undertaking, posttest forms were not always returned by the deadlines.

Videotaping

Pretest. Pretest videotape data collection was conducted over a four-week period beginning on February 12th, when student teachers had been in the classrooms approximately two weeks. One school week was spent taping at each of the four sites. Videotapes were made in the following site order: School 3 (week of February 12th), School 4 (week of February 19th), School 1 (week of February 26th), and School 2 (week of March 5th).

Several weeks prior to the scheduled dates for videotaping lessons at each school, copies of "Guidelines for Student Teachers and Public School Supervising Teachers Concerning Videotaping" were distributed. A cover letter to the public school supervising teacher requested that he or she keep one copy of the guidelines, give one copy to the student teacher, and select for the student teacher a group of 10 pupils, according to the instructions in the guidelines. The supervising teacher was further asked to select any unassigned videotaping time period convenient for him on the schedule posted in his school.

The guidelines (reproduced in their entirety in Appendix A) stressed that the lesson videotaping was in no way a test, and was not designed to show each student teacher at her best, but rather was intended to capture a representative sample of her teaching. Subjects were asked to teach the lesson as they would any other.

Subjects were instructed to write and center their teaching on one to three learner objectives relative to the following general goal:

> The student teacher will introduce a set of terms (e.g., concepts, symbols, technical terms...) which her pupils probably have not encountered previously. The terms may be drawn from the subject areas of Language Arts, Reading, Social Studies, or Science. The set of terms introduced should become part of the working knowledge of each learner.

Subjects were asked to plan for and teach a 20-minute lesson. It was suggested that they spend an hour or less in preparing the lesson. The guidelines further stated that subjects could use any materials they wished to bring to the videotaping room, and that they would be allowed to arrange the pupils and the room as desired, given the limitations of camera placement.

Another section of the guidelines provided instructions for the random selection of the 10 pupils to be sent for the videotaping.

The mechanics of the videotape lesson were described as follows:

> One member of the videotape crew will take you smoothly and naturally through the following steps. You will:
> 1. Bring your pupils to the room a few minutes early.
> 2. Arrange the pupils and room.
> 3. Announce to the cameraman that you are ready to begin. (A microphone will be placed around your neck at this time.)
> 4. Teach your lesson for 20 minutes (clock provided), announcing to the cameramen when you are through. (Because of the tight schedule, at exactly 20 minutes the videotape camera will stop, the crew will indicate this, and you will have a couple of minutes to end the lesson naturally. Then, if you are still not finished, the crew will have to interrupt you.)

5. Give your pupils' attention to the crew member in charge.
6. Complete a brief (2 minutes) questionnaire on how natural and representative you felt your teaching was, while the pupils do the same.
7. Return your pupils.

Subjects were asked to complete a one-page form summarizing the planned lesson, to be handed to the videotape crew at the time of the taping session. The student teachers were requested to call the person in charge of videotaping (name and phone number were given) if they had questions. Finally, a diagram of the floor plan student teachers would find in the portable buildings used for taping at all sites other than School 3 was included with the guidelines. At School 3, a regular classroom which had been converted to a student teacher lounge was used for videotaping.

Approximately one week after the guidelines were distributed, and one week before the start of videotaping, videotape scheduling "sign-up" sheets were made available in each school. The college supervisors co-ordinated this scheduling in their respective schools. The schedule was treated in a fairly flexible manner. It was not an infrequent occurrence for a subject to appear at her scheduled time only to reschedule the taping for a later time-slot.

Two evaluators trained and practiced with the videotape crew with the goal of creating a standard procedure wherein the presence of the videotape crew and evaluators would be as unobtrusive as possible. During the first two days of taping, both evaluators were present. As it was then apparent that the procedures were working satisfactorily, only one of the evaluators supervised subsequent sessions. The second evaluator served only as a "back-up" in case of need. This second eval-

uator was never used, however, after the first two days. The evaluator present was responsible for guiding the subject through the steps outlined in the guidelines. In addition, the evaluator wrote the ID number of each subject on the chalkboard for the camera to record as identification, gave timing cues when necessary, and administered the posttaping questionnaires to student teachers and pupils.

At the conclusion of the taped lesson, the evaluator, as smoothly as possible, handed the student teacher copies of "My Feelings During the Videotaping" and "Student Teacher's Evaluation of Videotape Session." The evaluator explained that directions were on each form, and asked the subject to be seated and to complete the two forms while her students completed a questionnaire.

The evaluator then got the attention of the pupils and proceeded with the administration of "Our Lesson." Procedures here varied somewhat due to the age range of the pupils and the necessity for attaining rapport with the pupils. In all cases, however, the evaluator explained that she wanted to find out what <u>each one</u> of the pupils thought about the lesson their teacher had just taught them. She stressed that what was wanted was each student's "very own opinion," and that his answer might be different from his neighbor's answer. The pupils were asked to mark an *X* over the face which showed how they felt about each statement. The evaluator gave one or more examples until she felt sure the pupils understood, and then passed out the questionnaires and proceeded. Each statement was read aloud at least once (more times for some of the primary students), and the evaluator paused until all pupils had responded. For the youngest groups of pupils, further guidance was pro-

vided on the first few items. For example, the evaluator would read the first item statement--e.g., "Miss Smith's lesson was very interesting." Then she might go on to comment, "Now, was the lesson very interesting to you? If you think it was, make an X over the smiling face. If you weren't interested in the lesson, make an X over the frowning face. If you just don't know, make an X over the 'Don't know' face, the middle face with the straight mouth."

After pupils had completed the form, the evaluator collected both teacher and pupil forms, thanked all the participants, and excused them.

The in-session procedures adopted for the videotape sessions seemed to work quite smoothly, and most subjects were extremely cooperative. The subjective impression of the evaluator present at all of the videotape sessions, however, was that the lesson guidelines were generally ignored at all schools other than School 2. In her opinion, only at School 2 had most of the subjects made a conscientious effort to teach a lesson which consisted of the introduction of a new set of terms or concepts. At the other three sites, although no formal count was made, the evaluator felt that many or most of the student teachers presented review rather than new material, which in many cases also did not involve a set of terms or concepts.

It was a commonplace for a much larger or smaller group of pupils than requested to appear for the session. In a number of cases, individual pupils attended several taping sessions, indicating that either 1) randomization procedures as outlined in the guidelines were not always followed, or 2) rosters of pupils were not cross-checked by the teachers as requested.

Posttest. Posttest videotape data collection was conducted over a four-week period ending one week before the end of the spring semester. Videotapes were made in the following site order: School 3 (April 16-18), School 1 (April 24-27), School 4 (April 30-May 2), and School 2 (May 3, 4, 7, and 8).

The procedures followed were generally the same as the pretest procedures. "Guidelines for Student Teachers and Public School Supervising Teachers for Second Videotaping Session" were distributed approximately two weeks prior to taping at each site, as before. These guidelines for the second taping (reproduced in Appendix A) were a streamlined version of the first set of guidelines. These instructions did request that students bring only 8 (rather than 10) pupils to the session, and specifically requested that review lessons not be taught. The evaluator with the videotape crew felt that there was more general compliance with these Guideline requests than in the pretest tapings.

Analysis of data from the pretest tapings indicated that continued use of "Our Lesson" for posttest data collection would not be worthwhile, due to a highly negatively skewed response distribution. The instrument was therefore dropped, and the evaluator's post-tape responsibility was limited to giving the student teacher "My Feelings During Videotaping" and a slightly revised version of "Student Teacher's Evaluation of Videotape Session." The evaluator chatted with the pupils and otherwise kept them orderly while the student teacher completed the two forms.

In addition, after each lesson was completed, the evaluator and the two videotape crew members independently completed "Our Lesson (Observer Rating)."

Chapter 5

DATA ANALYSIS STRATEGIES

There are two results of interest in the present investigation--main effects of program and trait-treatment interactions involving program. A main effect occurred when one program (PTE or non-PTE) produced a significantly higher ($p < .10$) mean on a criterion variable than the other program regardless of the personality traits and attitudes of the individual subjects. A trait-treatment interaction occurred when the programs produced differential criterion performance for different levels of personality and attitude held by the individual subjects.

Program main effects were detected with one-factor, two-level analyses of variance, separate analyses being performed for each of the 120 criterion variables employed in this study. Trait-treatment interactions were analyzed with standard statistical techniques which employed both the homogeneity of group regressions test (Edwards, 1968) and the test for regions of significance (Johnson and Neyman, 1936). This general methodology is described elsewhere in detail (Borich, 1971; Borich, 1972; Borich and Wunderlich, 1973). The trait variables employed were the 22 scale scores taken from the Comprehensive Personalization Assessment Battery, the Adjective Self Description, the Self Report Inventory, and the One Word Sentence Completion. The treatment variable was program (PTE versus non-PTE). One hundred and eleven criterion variables were selected for

trait-treatment interaction analysis. For most of these criterion variables, 22 interaction analyses were performed--one for each of the 22 trait variables, resulting in an unusually large number of trait-treatment interaction analyses. Impetus was therefore provided for the use and development of special data reduction procedures described below.

The basic strategy of the present evaluation was to sample a wide range of criterion variables. This strategy was responsible for both a large amount of data collection and a large number of analyses.

The goal of any data analysis scheme is to reduce a mass of data to a manageable number of findings or relationships. Normally, adequate data reduction is achieved through the use of standard statistical analyses (e.g., means, correlations, F-tests, etc.). The results of these analyses are then separately listed and discussed. In the present investigation, this straightforward procedure of separately listing all results is ruled out by the large number of analyses which had to be performed. Clearly, special data analyses and reporting procedures were called for. Attention is now shifted to these special procedures.

Data Analysis Strategy

Data analysis proceeded according to the following steps: data selection; statistical analysis of main effects and trait-treatment interactions; identification of chance results; and combination of results having the same psychological or educational meaning. Each of these steps will be considered in turn--the format being a general discussion followed by technical considerations.

Data Selection--General Discussion

Recall that criterion data were collected at two separate times: 1) midway through the training program and 2) at the end of the training program. Some variables were measured only at one time or the other. Other variables were measured at both times. For variables measured at two times (pre and post), attention in the present report has been largely limited to post data; and a reference to such variables is a reference to the post data, unless otherwise indicated. One reason for focusing on the post data is that these data, collected after full exposure to a training program, are the data most appropriate to demonstrate differences between the two training programs. A second reason was the similar patterns of results for the pre and post data.

When dealing with a large mass of data, it is important to reduce redundant information. Factor analysis is a technique well suited to that end. When conceived of as a data reduction technique, factor analysis can be viewed as reducing a given set of variables to a smaller set for which redundancy is minimized. In the present investigation, the following strategy was used with regard to factor analysis. 1) If previous research had established a factor structure for a data collection instrument, then that structure was usually accepted for the present purposes. 2) If an instrument was comprised of only a few items (six or less), then factor analysis was deemed unnecessary. 3) For instruments with more than a few items and with no previously established factor structures, factor analyses were computed.

---

*Pre data in this study refers to criterion data collected at the end of the first semester of the PTE program and represents about one half of the usual treatment. Post data refers to criterion data collected after the second semester of treatment, which includes the student teaching experience.

When a factor structure was accepted for an instrument, then a set of appropriate factor scores was computed from the original item scores. These factor scores provided the data for all subsequent analyses--the analyzed variables corresponding to the factors. Note that the variables corresponding to factors have already been described in Chapter 3. Thus, the results of all factor analyses have already been reported and need not be repeated here. Finally, when a factor structure was not accepted for an instrument, then the data to be analyzed consisted of the item scores for that instrument. The variables analyzed then corresponded to the items.

Following all factor analyses, the set of variables (both factor based variables and item based variables) to be used in subsequent analyses was scrutinized with regard to variance. If the variance for a variable was equal to or approached zero, then that variable was dropped from subsequent analyses.

This concludes the general discussion of data selection. The following section presents the technical details relevant to data selection. The reader with little interest in technical details may wish to skip this section.

Data Selection--Technical Details

Emphasis on post data. Some of the criterion instruments used in the present evaluation were administered only once; others were administered at two separate times. The criterion instruments for which both pre and post data were available were the Adjective Self Description, the Self Report Inventory, the three videotape coding systems (IASTV2, FAIR, COS), the Student Evaluation of Teacher Training Program, the Readiness Assessment, and the Profile of Learning Priorities. The post data from all of these instruments were thoroughly analyzed. While no analyses were performed on the pre data for the Readiness Assessment and the Profile of Learning Priorities, several analyses were performed on the pre data for the other six instruments. For the Adjective Self Description, the Self Report Inventory, IASTV2, FAIR, COS, the pre data program

differences appeared to be quite similar to the post data program differences--with regard to both the number of significant effects and the variables which demonstrated these effects. Therefore, only the post data results for these instruments have been included in the present report.

Finally, an effort was made to statistically verify the conclusion that the pre and post program differences were usually quite similar. This verification was limited to a single instrument (the IASTV2) as a test case. A repeated measures analysis of variance was performed on each of the 35 variables from the IASTV2. The between variable was program (PTE versus non-PTE), and the within variable was time of administration (pre versus post). If the effect of program was different for pre than for post, then a significant program by time of administration interaction should have been found. This interaction proved significant at the .10 level in only four of the 35 analyses. Since 3.5 of these interactions were expected to be significant at .10 on the basis of chance alone, there results support the conclusion that the patterns of pre and post results were similar.

Factor analysis. For many of the data collection instruments used in the present investigation, previous research had already established factor structures. Veldman (1970) presented a factor structure for the Adjective Self Description; Bown and Veldman (1967) for the Self Report Inventory; Watkins (1973) for the Teacher Concerns Checklist; Wehling and Charters (1969) for the Teacher Beliefs instrument; and Haak, Kleiber, and Peck (1972) for the SET II instrument (evaluation of the student teacher by her pupils). These previously established factor stuctures were accepted with the exception of that for the SET II instrument. The present version of the SET II differs from that factor analyzed by Haak, et al., in that it incorporates one new item and one revised item. Thus it was necessary to perform a new factor analysis on the present version of the instrument.

In the present investigation, a factor analysis was performed in the manner suggested by Veldman (1967). Orthogonal factors were derived using first a principal components and then an axis rotation procedure. Successful factor analyses were performed for the SET II, the Instructional Domain section of the Profile of Learning Priorities, and the Teaching and Motivational Plans section of the Professional Plans and Affiliations Questionnaire. The results of these three factor analyses have already been presented in Chapter 3.

An attempt was also made to factor analyze the three videotape coding systems (IASTV2, FAIR, COS) as a unit. Since the same videotaped lessons were coded with each of the three systems, it was hoped that a factor analysis would group together similar variables from the three systems. For example, it was hoped that the following variables would be grouped together: "Teacher Presentation" from the COS, "Teacher Lectures" from the FAIR, and "Lecturing--Giving New Information" from the IASTV2. However, the factors derived did not lend themselves to easy interpretation, and the factor structure for the three videotape coding systems was rejected.

A possible reason for the failure of this factor analysis was the small number of subjects relative to the number of variables. There were 76 variables and fewer than 77 subjects (after missing data was considered)--a condition which precludes the finding of a reliable factor structure (Fruchter, 1954). No further attempt was made to factor analyze the videotape coding systems, and the individual variables from these systems were accepted as criterion variables.

Two of the data collection instruments used in the present investigation were composed of only six items each. These instruments were the Readiness Assessment and Our Lesson. Because of the small number of items, factor analysis was deemed unnecessary, and the individual items from these instruments were accepted as criterion variables.

For two instruments--the Student Evaluation of Teacher Training Program and the One Word Sentence Completion--scales were derived without the use of formal factor analysis techniques. The Student Evaluation of Teacher Training Program instrument was constructed to measure the extent to which a teacher training program fulfilled ten objectives presented in the Basic Program Plan for the PTE program. Four to six items were generated for each of the objectives, the total number of items being 48. The scale for an objective simply consisted of the items generated for that objective. For the One Word Sentence Completion instrument, the scales used in the present investigation were those derived by Veldman (1971).

Computation of factor scores and scale scores. Simple unit-weighted factor scores were computed for the present purposes according to the following steps. First, each variable (or item) was assigned to the factor for which it had the loading of greatest absolute magnitude. For example, if variable X loaded +.65, -.76, and +.34 on factors I, II, and III, respectively, then variable X was assigned to factor II. As a result, nonoverlapping subsets of the variables from an instrument were assigned to the factors for that instrument. Second, when a variable assigned to a factor loaded negatively on that factor, then the scoring of that variable was reversed. If variable X--based on a 1-5 Likert scaled item--was to be reversed, then scores of 5 were changed to 1, and scores of 4 were changed to 2. Third, the factor score for a particular subject was then calculated as the simple sum of the relevant variable scores.

The simple unit-weighted factor measures employed in the present investigation differ from the more complex orthogonal regression weighted factor score. An orthogonal regression weighted score for a factor is calculated by summing the weighted variable scores for the entire set of variables entering into the factor analysis after each variable score has been weighted by the loading of that variable on the factor. The simple unit-weighted measure was chosen for the present purposes because it possesses several advantages over the more traditional measure. Among the advantages are ease of computation, less susceptibility to distortion in application to a new sample of subjects, and more direct interpretation (Schweiker, 1967). In addition, Veldman and Parker (1968) reported an instance in which the simple unit-weighted measure has greater external validity.

When scales were derived without the use of factor analysis (e.g., the Student Evaluation of Teacher Training Program instrument), then scale scores were calculated by summing the variable scores for the variables assigned to a scale.

Discarding of variables with low variance. Subsequent to the calculation of factor and scale scores, the criterion variables to be used in further analyses were examined with regard to variance. One trait variable from the One Word Sentence Completion was found to have no variance, and this variable (Rejection of a Teaching Career) was discarded. All of the subjects who had data for this variable had a score of 0--i.e., zero responses indicating rejection of a teaching career. Several criterion variables from the IASTV2 and FAIR videotape scoring systems were also found to have restricted variances. Recall that scores on videotape variables were in terms of the percentage of time that a particular behavior occurred. Videotape variables with mean scores of less than .1 across all subjects were discarded. That is, discarded variables were those which referred to exceedingly rare behaviors which occurred less than .1% of the time. Note that the variance for a variable with a mean of .1 is necessarily restricted. Using the above criterion for discarding videotape variables, seven IASTV2 variables, eight FAIR variables, and no COS variables were discarded. The variables discarded from the IASTV2 were as follows:

Variable No.	Name
1	Accept Feelings (Empathy)
11	Give Substantive Information--Previous Information
13	Justification of Authority
14	Controlled Silence--Demonstration
16	Controlled Silence--Looking at Notes
24	Student Questions--Procedural Open
26	Affective Response--Negative

Variables discarded from the FAIR were as follows:

Variable No.	Name
1	Teacher--Values (Rapport)
8	Teacher--Criticizes
9	Teacher--Yea (For Self)
11	Teacher--Owns Up

Variable No.	Name
16	Student--Encourages
17	Student--O.K.
22	Student--Resists
24	Student--Wool Gathering

Statistical Analyses for Program Main Effects

Program main effects were assessed for each of 120 criterion variables. For each criterion variable, simple analysis of variance was used to determine the statistical significance of the difference between the PTE group mean and the non-PTE group mean. A significance level of .10 was chosen for rejecting the null hypothesis of no program differences. The following technical details section on main effects analyses can be skipped without loss of continuity.

Technical details. All main effect analyses were performed with computer programs written by Veldman (1967). The number of subjects included in an analysis varied because of missing data. In the following discussion, let N_1 be the number of PTE subjects included in an analysis and N_2 be the number of non-PTE subjects included in that analysis.

For pre and post data analyses, a linear regression technique was used. The full model for the regression analysis was:

$$Y = a_0 + a_1X_1 + a_2X_2 + e_1 \qquad (1)$$

where Y = the criterion variable,
a_0 = the regression constant,
a_1 = the regression coefficient (b-weight) for X_1,
X_1 = a dichotomous predictor variable equaling 1 for PTE subjects and 0 for non-PTE subjects,
a_2 = the regression coefficient (b-weight) for X_2,
X_2 = a dichotomous predictor variable equaling 1 for non-PTE subjects and 0 for PTE subjects,
e_1 = the residual error.

Note that there is only one independent predictor variable in equation (1), since there is a perfect negative correlation between X_1 and X_2. The restricted model for the regression analysis was:

$$Y = a_0 + e_2 \qquad (2)$$

where Y = the criterion variable,
a_0 = the regression constant,
e_2 = the residual error.

An iterative technique was employed to obtain least squares estimates of the parameters a_0 and a_1. Since the information contained in variable X_2 was completely redundant to that contained in variable X_1, it was unnecessary to estimate the value of the parameter a_2--the value of a_2 always being equal to 0. The sum of squares (SSe_1) based upon the e_1 residual errors and the sum of squares (SSe_2) based upon the e_2 residual errors were then used to calculate the following F ratio:

$$F = \frac{(SSe_2 - SSe_1)/1}{SSe_1/(N_1 + N_2 - 2)} \qquad (3)$$

The degrees of freedom for this F are 1 for the numerator (i.e., the number of independent parameters in the full model minus the number of independent parameters in the restricted model) and $N_1 + N_2 - 2$ for the denominator (the total number of scores minus the number of independent parameters estimated for the full model). An approximation technique was employed to estimate the significance level of the obtained F ratio (Veldman, 1967).

Statistical Analyses of Trait-Treatment Interactions

General comments. Trait-treatment interactions involve the relationships between entering competencies, personality traits, and attitudes with hypothesized program outcomes. To assess trait-treatment interactions, trait-criterion relationships are examined separately for each treatment (program) in order to determine the extent to which these relationships differ across treatments. As trait-criterion relationships differ from one treatment to another, the likelihood of a trait-treatment interaction increases. For example, if the relationship between a particular trait and the criterion is positive for one treatment and negative for a second, students who score low on the trait may perform better when assigned to the second treatment, and students who score high on the trait may perform better when assigned to the first treatment. Such findings allow the differential assignment of students to programs based upon their entering

competencies, personalities, and attitudes.

The development of a methodology for studying trait-treatment interactions has been in recent years a primary concern of the differential psychologist. This concern has stimulated the development and rejuvenation of several specialized statistical techniques to describe the effect of an interaction in terms of student traits which can be used to assign students to the treatment for which they are best suited. Such a methodology differs from traditional factorial designs in that variables commonly dichotomized or trichotomized to fit the requirements of analysis of variance are not divided into discrete categories with a resultant loss of information and an increase in error of measurement but rather are used in their continuous form to describe as many different types of students as there are values of a particular trait. In trait-treatment interaction designs, discrete groups (e.g., high, average, low on trait X) are replaced with the observed continuum of values that represents the entering competencies, personality traits, and attitudes of each student in the instructional program.

While the actual calculation of a trait-treatment interaction is a complex statistical procedure, the general methodology used in the present investigation may be summarized in two steps. The first step was to calculate regression lines for each of the trait-criterion relationships and to test the homogeneity of regression slopes--i.e., the extent to which these slopes differ across treatments. When regression slopes were found to differ at the .10 level of significance, a second step was employed to determine the region(s) of trait values for which

the treatments were significantly different. These regions of significance were determined with the Johnson-Neyman technique (Johnson and Neyman, 1936). When regions of significance were obtained, the percentage of subjects falling within those regions was calculated. This percentage was interpreted as a measure of the efficiency with which the trait variable can be used to assign subjects to treatments.

Consider the following hypothetical result. Pictured in Fig. 5-1

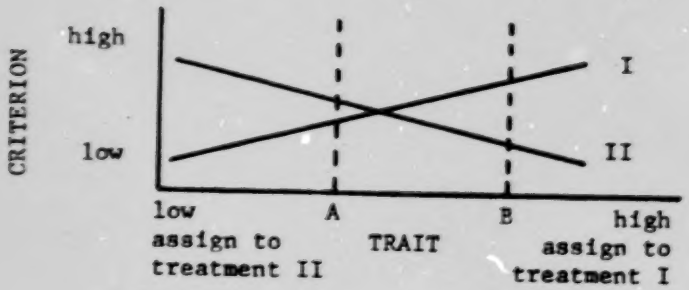


FIG. 5-1. Example of trait-treatment interaction. Two different treatments are represented by I and II.

are significantly different regression slopes for two treatments with a region of significance to the left of point A and to the right of point B. Students with a trait score above value B are assigned to treatment I, while students who scored below value A are assigned to treatment II. For students scoring between values A and B, both treatments are considered equally suitable for producing the criterion behavior.

Trait-treatment interactions analyzed. Twenty-two trait-treatment interactions were examined in the present investigation. These 22

interactions involved combinations of one treatment variable (program--PTE versus non-PTE) with each of 22 trait variables. These trait variables have been listed and described in Chapter 3. They represent data collected prior to the start of the teacher training program. Each of the 22 trait-treatment interactions was examined with regard to a large number of post data criterion variables (97 criterion variables for some interactions and 96 for others). The following section deals with the technical details of the trait-treatment interaction analyses. This section may be skipped without loss of continuity.

Technical details. The homogeneity of regression tests were performed with a computer program written by Veldman (1967). The Johnson-Neyman analyses for regions of significance were computed with a program written by Borich (1971). The number of subjects included in a specific analysis varied because of missing data. In the following discussion, let N_1 be the number of PTE subjects included in an analysis and N_2 be the number of non-PTE subjects included in that analysis.

The basic test for a significant trait-treatment interaction effect--the homogeneity of group regressions test--employs a full and a restricted linear regression model. The full model is:

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_3 + a_4X_2X_3 + e_1 \qquad (4)$$

where Y = the criterion variable,
a_0 = the regression constant,
a_1 = the regression coefficient (b-weight) for X_1,
X_1 = a dichotomous predictor variable equaling 1 for PTE subjects and 0 for non-PTE subjects,
a_2 = the regression coefficient (b-weight) for X_2,
X_2 = a dichotomous predictor variable equaling 1 for non-PTE subjects and 0 for PTE subjects,
a_3 = the regression coefficient (b-weight) for the X_1X_3 product variable,
X_3 = the trait variable,
a_4 = the regression coefficient (b-weight) for the X_2X_3 product variable,
e_1 = the residual error.

Note that there are only three independent predictor variables in equation (4), since there is a perfect negative correlation between X_1 and X_2. The restricted model for the analysis was:

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + e_2 \qquad (5)$$

where Y, a_0, a_1, X_1, a_2, X_2, and X_3 are defined the same as for equation (4). In equation (5), a_3 is the regression coefficient for X_3, and e_2 is the residual error.

An iterative procedure was employed to obtain least squares estimates of the parameters in equations (4) and (5). Four parameters (a_0, a_1, a_3, and a_4) were estimated for equation (4), and three parameters (a_0, a_1, and a_3) were estimated for equation (5). Since the information contained in variable X_2 was completely redundant to that contained in variable X_1, it was unnecessary to estimate the value of a_2 for either equation--the value of a_2 always being equal to 0. The sum of squares (SSe_1) based upon the e_1 residual errors and the sum of squares (SSe_2) based upon the e_2 residual errors were then used to calculate the following F ratio.

$$F = \frac{(SSe_2 - SSe_1)/1}{SSe_1/(N_1 + N_2 - 4)} \qquad (6)$$

The degrees of freedom for this F ratio were equal to 1 for the numerator (number of parameters estimated for the full model minus the number of parameters for the restricted model) and equal to $N_1 + N_2 - 4$ for the denominator (total number of scores minus the number of parameters estimated for the full model). The exact chance probability associated with this F ratio was estimated with an approximation technique (Veldman, 1967).

The chance probability associated with the F ratio given in (6) is the probability that homogeneity of regression slopes exists. When this probability was below .10, the null hypothesis of homogeneity of regression was rejected, and a significant trait-treatment interaction was assumed to exist.

The Johnson-Neyman technique (Johnson and Neyman, 1936) was applied to significant trait-treatment interactions. This technique involves the following procedures. Regression lines for the criterion variable on the trait variable are calculated separately for the two programs. Let Y_1' be the predicted criterion score for the PTE group and Y_2' be the predicted criterion score for the non-PTE group. Note that different values of Y_1' and Y_2' will be obtained for each different value of the trait variable. (See Figure 1.) Likewise, the difference between Y_1' and Y_2' will vary for different values of the trait variable. The strategy of the Johnson-Neyman technique is to determine the trait variable values for which the $Y_1' - Y_2'$ difference is significant. A t statistic is computed in the following manner:

$$t = \frac{Y_1' - Y_2'}{S_D} \qquad (7)$$

where S_D is an estimate of the standard error of the difference between predicted scores. The degrees of freedom associated with this t statistic are equal to the number of subjects minus 4 (i.e., $N_1 + N_2 - 4$ in our case).

Four degrees of freedom are lost because four parameters are estimated--a b-weight and a regression constant for each of the two programs (PTE and non-PTE). In the present investigation, the t value required for significance was that with a chance probability of .05.

Now a different t value can be calculated for each different value of the trait variable. When the t is significant, then the trait value for which that t was calculated falls within a region of significant differences between the two programs. If a t is not significant, then the corresponding trait value does not fall within a region of significance.

The identification of regions of significance would be a simple matter if it were not for the fact that S_D as well as $Y_1' - Y_2'$ varies for different values of the trait variable. The value of S_D is a positive function of the magnitude of the $Y_1' - Y_2'$ difference and a positive function of the magnitude of the difference between the trait value in question and the mean of the trait variable (Walker and Lev, 1953, p. 400). The fact that both $Y_1' - Y_2'$ and S_D are varying with the value of the trait variable makes the situation complicated, and several different types of results are possible. Figure 5-2 presents the different results that can be found with the Johnson-Neyman technique. In Figure 5-2, regions of significance are shaded. Consider the result in 5-2(a). Here there are no regions of significance. In Figure 5-2(b), two regions of significance are present. The lower or left region of significance has an upper bound at point A and no lower bound. The upper or right region of significance has a lower bound at point B and no upper bound. It is not necessary that these two regions of significance be symmetric about the point of intersection. In Figure 5-2(c), there is only one region of significance. This region falls below the point of intersection of the two regression lines and has a lower bound of point A and an upper bound of point B.

Figure 5-2(c) represents a quite unusual result. There is a significant program difference at point A but no significant difference for points below A, even though the magnitude of the difference between programs (i.e., $Y_1' - Y_2'$) is increased. This condition occurs when the value of S_D increases at a faster rate than $Y_1' - Y_2'$, as we move away from the point of intersection. In the present investigation, results similar to those depicted in Figure 5-2(c) occur only when the difference between the two regression slopes is marginally significant.

Identification of Significant Results Due to Chance

One hundred and twenty analyses were performed for the detection of main effects. Given the chosen significance level of .10, 12 of these main effects analyses were expected to be significant on the basis of chance. This large number of expected chance results presents serious problems to interpretation. These problems are greatly magnified with

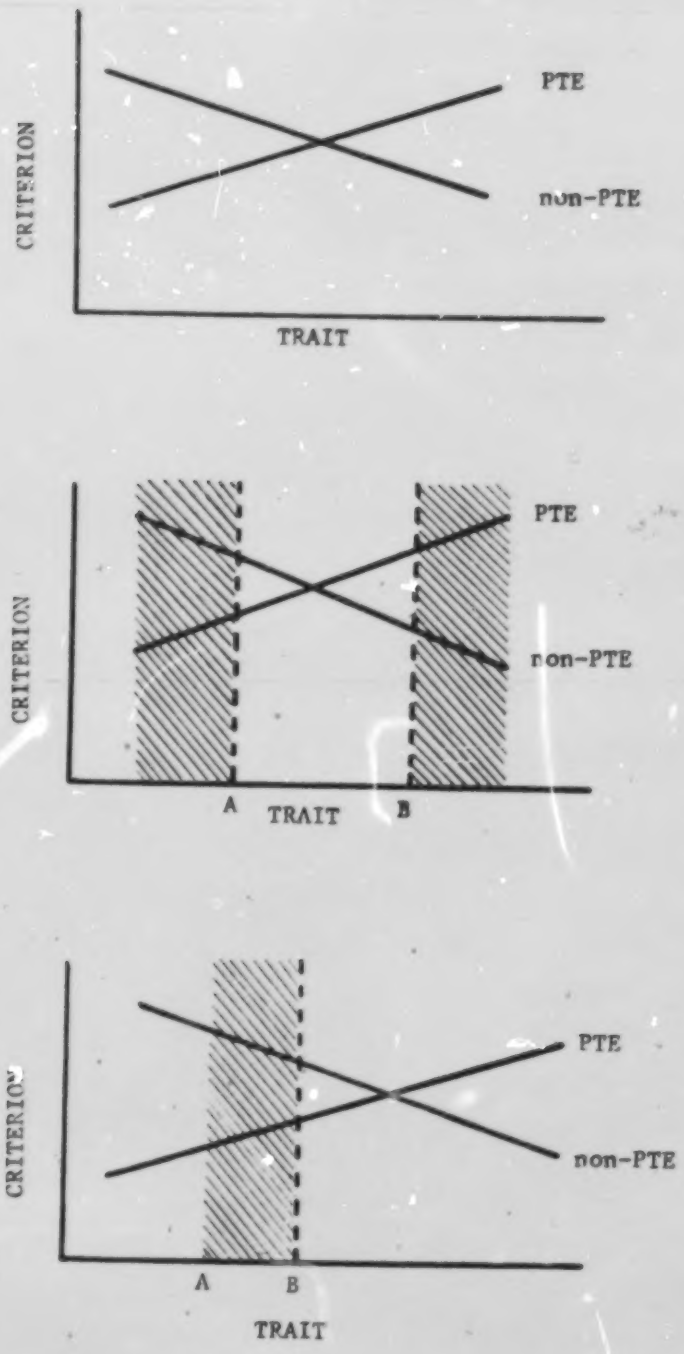


FIG. 5-2. Different results obtainable with the Johnson-Neyman technique. Shaded areas represent regions of significant differences between the PTE and non-PTE programs. (a) No regions. (b) A left and a right region. (The left region has no lower bound, and the right region has no upper bound.) (c) One region (both upper and lower bounds present).

regard to analyses of trait-treatment interactions, the number of such analyses exceeding 2000. Clearly, it would be of great value if we could detect if certain differences were due to chance and eliminate those differences from further consideration.

Because of the nature of inferential statistical analysis, an obtained difference associated with a small chance probability (p) cannot definitely be attributed to the independent variable. Such a result occurs by chance with probability p. As long as p is not 0, there is still a chance that the results are spurious. The standard method for handling this dilemma is to set an arbitrary significance level and then to accept as valid all results which surpass this level.

There are certain cases in which a significant difference can be discounted. Consider the case of an experiment which produces a significant result (at the .10 level) but, in nine subsequent replications, no significant results are obtained. Certainly our confidence in the initial result is shaken. The obvious inference would be to attribute this initial significant difference to chance. In contrast, if several of the replications had produced significant differences similar to the initial difference, then our confidence in the initial result would have been strengthened.

Now consider the case in which we have ten different criterion variables all measured in the same investigation. Interest is focused on whether two groups of subjects differ with regard to each of these variables. This situation is quite similar to a replicated experiment. For the purposes of this study the same logic will be applied, the

analyses being treated as if they were replicated experiments. For example, if the difference between the two groups is significant (at the .10 level) with regard to one of the criterion variables but not the nine others, we would conclude that we are dealing with a chance effect. As the number of obtained significant differences increases, the hypothesis that the results are produced entirely by chance becomes more and more questionable.

In the present investigation, such a chance-identification procedure was employed. A set of analyses was chosen and then the results for those analyses were examined. If the number of significant differences did not exceed the number expected on the basis of chance, then the entire set of results was taken as being nonsignificant and was excluded from further consideration. This procedure was applied separately to main effects analyses and to trait-treatment interaction analyses.

For main effects analyses, the first step was to check to see if the total number of obtained significant main effects exceeded the number expected on the basis of chance. When the obtained number was found to exceed the chance expected number, then each criterion instrument was examined individually. The total number of significant results for all the variables relevant to a criterion instrument was compared to the chance expectation. If the number of obtained significant results did not exceed the number expected by chance, then the significant results were rejected. If the obtained number exceeded the chance number, then the significant results were accepted as valid. For trait-treatment interaction analyses, a similar chance-identification procedure was employed. The total number of analyses

was considered and then each of the criterion instruments was examined in turn.

A few general comments about the chance-identification procedure should be made. First, the pattern of results was not examined prior to the determination of the chance-identification scheme. Thus there should have been no bias toward selecting favorable results. Second, the chance-identification procedure was quite conservative in nature. The confidence which can be placed in the remaining significant results far exceeds that associated with the chosen significance level of .10. That is, interest has been limited to relatively robust effects. Third, the results discussed in the present report are only those surviving the chance-identification test.

The specifics of the chance-identification procedure are discussed in the following technical section. This section may be skipped without loss of continuity.

Technical details. Consider a set of N analyses, each comparing two groups. Assume that the populations from which the groups are drawn do not differ--i.e., there is no "real" difference between the two groups. Any obtained difference will be entirely due to sampling error--i.e., to chance. Each of the N analyses can be conceived of as an independent event with two possible outcomes. The first possible outcome is the lack of a significant group difference at the .10 level, and the probability for the occurrence of this outcome is .90. The second possible outcome is a significant group difference, and the probability of this outcome is .10. The binomial expansion is applicable to this situation, in which we have N independent events with fixed-probability binary outcomes.

Say that a number (r) of the N analyses prove to be significant. Then we can use the binomial expansion to determine the chance probability of obtaining exactly r significant results and N - r non-significant results. The probability (P_r) of exactly r significant results can be calculated as follows:

$$P_r = \frac{N!}{r!(N - r)!} p^r q^{N - r} \qquad (8)$$

where p is the probability of a significant result (i.e., .10), and q is the probability of a nonsignificant result (i.e., 1 - p or .90). In similar fashion, we can determine the chance probabilities for r + 1, r + 2, r + 3, ..., r + (N - r - 2), r + (N - r - 1), and N significant results. Adding together the chance probabilities for r, r + 1, r + 2, r + 3,..., r + (N - r - 2), r + (N - r - 1), and N significant results, we obtain the chance probability for at least r significant results. Now if this chance probability is large (say greater than .10), then the entire set of results can easily be attributed to chance. However, if the chance probability of at least r significant results is small (say less than .10), then it is more difficult to discount the obtained significant differences as being due to chance. In this latter case, we reject the chance basis for the significant results. Obtaining at least that many significant results on the basis of chance is a relatively rare occurrence.

As an aside, note that what is involved here is really a second-order significance test. We are testing the significance of the number of obtained significant results. Initially we test each criterion variable for significance (at the .10 level). Then we examine the set of results for several criterion variables. If the number of significant results in this set of results is itself significant (at the .10 level), then we accept the individual significant results. If the number of significant results is not itself significant, then we discount the individual significant results. The resulting significance level for acceptance of a result as not being chance-produced has been substantially lowered--being closer to .01 than the original .10.

When N (the number of analyses) is large, then a normal curve approximation to the binomial distribution can be employed for determining second-order significance. This approximation is given by McNemar (1966, p. 61) as:

$$z = \frac{\text{obtained proportion} - \text{expected proportion}}{\left(\frac{pq}{N}\right)^{1/2}} \qquad (9)$$

where p and q are the probabilities associated with the two outcomes, and N is the number of events. For our purposes, p is the probability of significance (.10), and q is the probability for nonsignificance (1 - p or .90). If r significant differences are obtained, then the obtained proportion is r/N. The expected proportion is always pN. When r significant results are obtained, then a corresponding z value (normal deviate) can be calculated. The probability of at least r significant results is then approximated by the probability of obtaining a z at least as large--this latter probability being obtainable from a normal curve table.

McNemar's suggestions were followed with regard to the application of this normal curve approximation. First, equation (9) was applied only when the expected number of significant results was greater than 10--i.e., when N was greater than 100. Second, when the expected number of significant results was between 5 and 10 (N between 50 and 100), a correction for continuity was incorporated into equation (9). This correction for continuity involved subtracting .5/N from the numerator in equation (9). Third, when the expected number of significant results was below 5 (N less than 50), the normal curve approximation was not used.

When N was less than 50, the chance probability for the number of significant results could have been directly calculated using the binomial expansion. However, this tack was rejected for the present purposes. To understand why, consider what occurs when N is very small. If N is equal to 2, then the chance probability for obtaining at least one significant result is .19, and the chance probability for obtaining two significant results is .01. That is, when N equals 2, we must find two significant results in order to achieve second-order significance. Second-order significance is not achieved when only one result is significant--even though this represents one-half of the results. Similar problems can be demonstrated when N equals 3, 4, or 5. In these cases, a relatively high proportion of the analyses can be significant without this proportion itself achieving significance.

While the binomial expansion does not provide satisfactory results when N is very small, it is clear that for a moderately large N (say 20) the binomial distribution should prove satisfactory. But what is the cutoff point below which binomial expansion calculations should be avoided? Rather than guess at this cutoff, it was decided to use the binomial expansion only when N was 50 or more--i.e., when the normal curve approximation could be used. A less efficient, simpler method of identifying chance results was employed when N was less than 50. This method consisted of a simple comparison of the obtained number of significant results (r) with the expected number of significant results (.10N). When r exceeded .10N, then all significant results were accepted as valid. When r was equal to or less than .10N, then all significant results were identified as chance results.

Identification of chance results proceeded in the following manner for main effects. The second-order significance (.10 level) of the total number of obtained significant main effects was tested using the normal curve technique. Then the set of analyses for each criterion instrument was considered. N was always less than 50, and the simple comparison method was employed for identifying chance results.

The identification of chance results for the trait-treatment interaction analyses was more complicated. The following steps were followed. 1) The second-order significance (.10 level) of the total number of significant results was tested using the normal curve approximation. 2) For each criterion instrument, the second-order significance (.10 level) of the number of significant results for that instrument was tested. The number of analyses relevant to a criterion instrument was always greater than 50, and the normal curve approximation was used to test second-order significance. If second-order significance was not obtained, then all significant results for that criterion instrument were discarded. 3) If second-order significance did exist for a criterion instrument, then the results for each variable included in that instrument were examined. There were 22 trait-treatment interaction analyses for each criterion variable--one for each of the trait variables. The simple comparison method was used to identify chance results with regard to a criterion variable. If the obtained number of significant results did not exceed the 2.2 results expected on the basis of chance, then any significant results for that criterion

variable were discarded. 4) Finally, the remaining significant trait-treatment interaction results were catalogued with regard to the trait variable involved. If no more than two significant results were found to be associated with a trait variable, then that trait variable, as well as any significant results associated with it, was discarded. This final step was intended to preclude discussion of relatively impotent trait variables.

This concludes discussion of the data analysis strategy employed for the present report. Interest is now shifted to the actual results obtained.

Chapter 6

MAIN EFFECT COMPARISONS BETWEEN THE TRADITIONAL AND PERSONALIZED MODELS OF TEACHER TRAINING

Of a total of 120 analyses for program main effects, 24 analyses showed significant differences (.10 level) between PTE and non-PTE student teachers. The chance probability of obtaining 24 or more significant results is quite small ($z = 3.65$; $p = .0002$), and the total set of significant main effects was identified as not being chance-produced. The 24 criterion variables which produced significant main effects, as well as the characteristics of those effects, are presented in Table 6-1.

Further attempts to identify chance-produced significant results involved consideration of the results, criterion instrument by criterion instrument. The number of significant results for a criterion instrument was compared to the number of significant results expected by chance for that instrument. If the obtained number of significant main effects did not exceed the expected number, then significant results were identified as due to chance. If the obtained number of significant differences exceeded the expected number, then significant results were retained as valid. Only one of the 24 significant main effects was identified as being chance-produced--this main effect being that associated with the Teacher Initiates variable from the FAIR observation system. Discussion will be limited to the remaining 23 significant results.

One further matter must be considered before our attention can be turned to the significant main effects obtained in the present investiga-

Table 6-1

Significant Main Effects

Instrument Name	Scale Name	$\bar{X}$ PTE	$\bar{X}$ Non-PTE	% Var.	F	P	df
Career Behavior	Career Behavior	145.10	109.00	5.71	4.55	.03	1/75
FAIR	Student Brings Out	19.19	9.41	11.48	6.87	.01	1/53
	Student Solitary Work	5.46	12.73	9.52	5.58	.02	1/53
	Teacher Initiates[a]	.85	.21	13.68	8.40	.005	1/53
IASTV2	Student Questions--Procedural Closed	.65	.98	5.53	3.10	.08	1/53
	Overt Student Activity	5.67	9.92	5.63	3.16	.08	1/53
	Student Questions--Substantive Open	2.10	.18	15.70	9.87	.003	1/53
	Student Statements--Open	7.00	1.32	9.50	5.56	.02	1/53
	Flexibility Ratio	8.16	7.38	6.53	3.70	.06	1/53
COS	Teacher Initiated Problem Solving	22.82	18.54	4.87	2.71	.10	1/53
	Convergent Evaluative Interaction	32.58	37.91	5.03	2.81	.10	1/53
Teacher Concerns Checklist	Factor 1: Personal Concerns	3.17	3.51	7.69	5.77	.02	1/69
	Factor 2: Professional Concerns	2.85	3.25	11.09	8.58	< .005	1/69
	Factor 3: Concern for Pupils	3.79	4.12	6.63	4.91	.03	1/69

Table 6-1 continued

Significant Main Effects

Instrument Name	Scale Name	X PTE	X Non-PTE	% Var.	F	P	df
Adjective Self Description	ASD 1: Attitude	35.21	33.64	5.13	3.73	.054	1/69
Student Evaluation of Teacher Training Program	Program Integration	15.33	13.41	7.87	5.38	.02	1/63
	Individualized Teaching	17.22	14.10	13.81	10.09	.002	1/63
	Behavior Modeling[b]	21.39	20.17	4.55	3.00	.08	1/63
	Negotiation	15.80	13.96	5.15	3.42	.06	1/63
	Teacher Preparation	20.22	18.00	7.98	5.47	.02	1/63
	Person-Centered[b]	19.39	16.34	19.18	14.95	.001	1/63
	Personal, Intellectual and Social Development[b]	19.58	16.00	24.76	20.73	.001	1/63
	Teacher Educator-Student Teacher Interaction	14.83	13.03	8.66	5.98	.02	1/63
Teacher Beliefs	Consideration of Student Viewpoint	40.09	38.08	5.73	4.132	.04	1/68

[a]This result was identified as due to chance.

[b]These results were discarded because significant trait-treatment interactions occurred for these variables.

tion. Statistical methodologists have frequently pointed out (e.g., Winer, 1962; McNemar, 1966) that it is incorrect to interpret a main effect when the variables involved are also involved in a significant interaction. Consider the following example. Say that we find a significant program main effect (PTE versus non-PTE) for criterion variable X, such that the PTE group is superior with regard to variable X. Also assume that a significant trait-treatment interaction involving program is found for criterion variable X, such that Figure 6-1 represents the state of affairs.

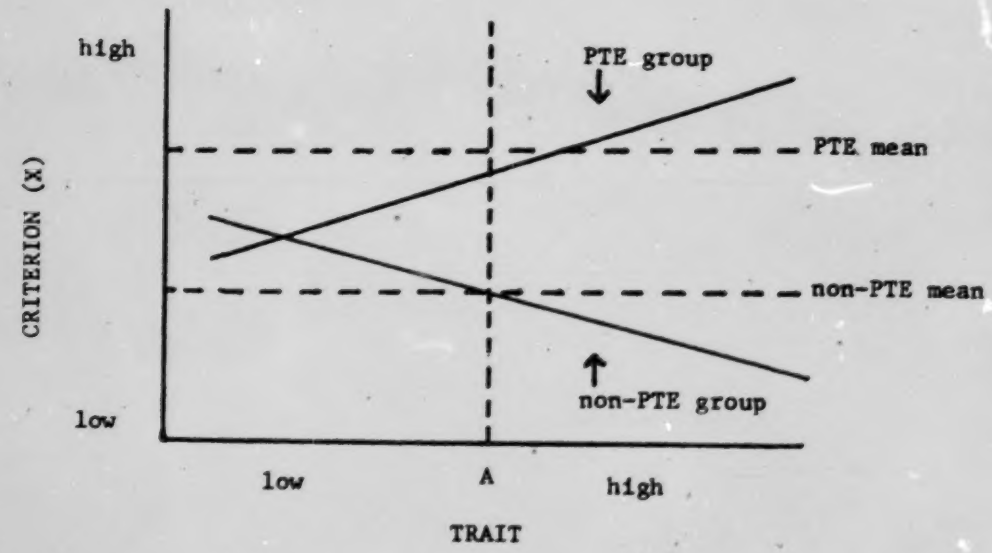


FIG. 6-1. Example trait-treatment interaction where program main effect is also present. Point A is the lower bound of a region of significant differences between the two groups.

Note in Figure 6-1 that the PTE mean is greater than that of the non-PTE group, this differnce being the basis for the obtained significant main effect. But further note that the PTE superiority is not general: the difference in means is produced mainly by individuals scoring high on the trait variable. Given this situation it would be incorrect to interpret

the significant main effect as implying a general superiority of the PTE group. Such an interpretation would greatly oversimplify and misrepresent the true situation. A more adequate interpretation of the situation can be based upon the significant trait-treatment interaction. This latter interpretation would establish 1) that the relationship between the trait variable and criterion variable is positive for the PTE group but negative for the non-PTE group and 2) that (based upon the application of the Johnson-Neyman technique) the groups differ significantly only for subjects who score above point A on the trait variable. Thus, the interpretation of the significant interaction provides an adequate picture of the true state of affairs. Interpretation of the main effect is unnecessary and unwarranted in this situation.

The present results were examined to determine if significant main effects and trait-treatment interactions occurred simultaneously for the same criterion variable. Such simultaneous occurrence precludes discussion of the main effect. Significant trait-treatment interactions were found for only three of the criterion variables demonstrating main effects. These three variables--1) Person-Centered; 2) Personal, Intellectual and Social Development; and 3) Behavior Modeling--all come from the Student Evaluation of Teacher Training Program instrument. Since the existence of trait-treatment interactions calls into question the generality of the main effects associated with these variables, these three significant main effects will not be considered. The following discussion of significant main effects has been limited to the 20 remaining significant main effects.

The significant main effects will be discussed below according to their classification in terms of the Type of Measurement X Domain of Competence Matrix described in Chapter 2 of this report. Each of the 120 criterion

variables analyzed for program main effects were classified according to this matrix. Each of the nine cells of the matrix will be considered in turn.

Cell 1: Self-Report X Intrapersonal Behaviors. Of the 11 main effect analyses for this cell, only one analysis detected significant program differences. On the Teacher Concerns Checklist, PTE students showed significantly less Personal Concerns ($p = .02$) than non-PTE student teachers. This significant effect for Personal Concerns may indicate that the PTE Program leads students to worry less about whether they are well-adjusted in these respects. However, in general the results for this cell are disappointing. The intrapersonal domain of competence was adequately sampled by the 11 self-report scales employed, and the psychometric properties of 8 of the 11 scales have been demonstrated to be adequate. This assurance of adequate measurement, coupled with the finding of only one program main effect for these 11 criteria, leads to the conclusion that the PTE Program does not generally enhance student teachers' feelings of intrapersonal adequacy or well being.

Cell 2: Self-Report X Interpersonal Behaviors. Three of the 11 analyses for this cell showed significant program main effects. For the Attitude variable from the Adjective Self Description instrument, the PTE student teachers were found to report more positive attitudes (also interpretable as greater social warmth) than non-PTE student teachers ($p = .05$). For the Consideration of Student Viewpoint variable from the Teacher Beliefs instrument, the PTE student teachers were found to report greater consideration of the viewpoints of their pupils than the non-PTE student teachers ($p = .04$). For the Concern for Pupils variable from the Teacher Concerns Checklist, PTE student teachers were found

to report less concern about their pupils than non-PTE student teachers ($p = .01$). This last effect may be interpreted as the PTE student teachers expressing less worry (negative concern) about their relationships with their pupils and their adequacy in meeting the needs of their pupils. The results for this cell then are quite encouraging, as they provide a consistent picture of PTE trainees being more considerate and warm with regard to their pupils and also being less worried about them.

Cell 3: Self Report X Career-Related Behaviors. Only one of 11 main effect analyses for this cell provided significant group differences, and the results for this cell are therefore somewhat discouraging. The sole significant effect occurred for the Professional Concerns variable from the Teacher Concerns Checklist. PTE trainees reported having more Professional Concerns than did non-PTE trainees ($p = .005$). This result implies, for example, that, relative to non-PTE student teachers, PTE student teachers worry more about whether the curriculum is appropriate, are more frustrated by the routine of the situation, and feel greater pressure.

Cell 4: Other Report X Intrapersonal Behaviors. No significant program main effects were detected in the five analyses for this cell.

Cell 5: Other Report X Interpersonal Behaviors. None of the six analyses for this cell demonstrated significant main effects.

Cell 6: Other Report X Career-Related Behaviors. No significant program main effects were detected in the four analyses for this cell.

Cell 7: Observation X Intrapersonal Behaviors. The present investigation did not attempt to study any variables which would be classified within this cell. Only physiological measures would provide criterion variables for this cell.

Cell 8: Observation X Interpersonal Behaviors. No significant program main effects were found for the five analyses relevant to this cell.

Cell 9: Observation X Career-Related Behaviors. Of the total of 20 significant program main effects, 10 were obtained from the series of 57 analyses performed on criterion variables classified as "Observation--Career Related." The only analyses in the cell based on data from a source other than the videotaped lessons showed a clear positive effect for the PTE Program. In this analysis, PTE student teachers were shown to have activated Teacher Placement files at the University of Texas Teacher Placement Service earlier than Controls. The score on this variable, called Career Behavior, was the number of days prior to June 1, 1973, that the student teacher activated a placement file. Since the date the student teacher enrolled in his practicum course might have influenced the date on which a placement file was opened, the time of enrollment in the practicum was considered as a rival causal variable. The correlation between date of enrollment in the practicum and date of activating the placement file failed to reach significance even with a liberal alpha level of .10. Date of enrollment in the practicum was thus ruled out as a rival causal variable, and it was concluded that student teachers in the PTE Program evidenced significantly stronger career-relevant behavior ($p = .03$) as indicated by the fact that PTE's, on the average, activated placement files over a month earlier than did Controls.

Interpretation of the videotape findings is somewhat problematic. Program main effects were found for 9 of the 56 videotape scales analyzed in this cell of the matrix. However, the validity of the videotaped lesson procedure as a mechanism for program evaluation may be somewhat questionable, since (1) the reliability of the coding on several of the scales is quite low,

and (2) some of the scales may have been inappropriate for a 20-minute sample of classroom interaction under the conditions which prevailed in this study. It should also be pointed out that, in the opinion of the evaluator who co-ordinated videotaping at all sites, some of the significant findings may have resulted from differential adherence to the guidelines for the videotaped lessons. It has been suggested that some of the PTE student teachers may have tended to deviate from the guidelines to a greater extent than Controls, by ignoring the request that new material be presented to the students during the lesson. However, no objective evidence supporting this contention (e.g., differences on the Give Substantive Information--Lecturing, New Information variable from the IASTV2; differences on the Teacher Presentation variable from the COS; and differences on the Teacher Lectures variable from the FAIR) was found.

Such difficulties are not unexpected with regard to observational systems employed in a research project evaluating on-going educational programs, and they certainly do not preclude the value of the present videotape results. However, such factors should be kept in mind when the implications of the videotape results are considered.

Of the nine significant main effects obtained for videotape variables, five occurred for variables from the IASTV2 coding system. The pupils taught by PTE trainees asked significantly more Substantive, Open Questions ($p = .003$) and made more Open Statements ($p = .02$) than the pupils of non-PTE trainees. On the other hand, pupils taught by non-PTE student teachers asked more Closed, Procedural Questions ($p = .08$) and engaged in more Overt Activity ($p = .08$) than pupils taught by PTE student teachers. Finally, PTE trainees demonstrated more Flexibility (higher Flexibility Ratio) by using more types of interactions with their pupils ($p = .06$). As pointed out in the IASTV2

manual (Hall, 1972), a high Flexibility Ratio can be either positive or negative, since a teacher could obtain a high score either through willingness to shift strategy when it becomes apparent that an initial approach is not working, or through a lack of focus on any consistent strategy.

Two significant main effects were found for the FAIR coding system. Pupils taught by PTE trainees Bring Out more ($p = .01$) than pupils taught by non-PTE trainees. On the other hand, pupils of non-PTE student teachers engaged in more Solitary Work ($p = .02$) than pupils of PTE student teachers. The remaining two significant main effects were associated with the COS videotape scoring system. Greater Teacher Initiated Problem Solving ($p = .10$) was found for PTE trainees. In contrast, a Convergent Evaluative Interaction ($p = .10$) pattern tended to occur more frequently in non-PTE classrooms.

These results clearly indicate that different teacher-pupil interactive patterns were occurring for the two programs. The non-PTE classroom fits the image of the conventional or traditional classroom. The locus of control is firmly with the teacher, and open-ended interaction between the pupils and teacher is minimized. Prominent pupil behaviors involve activities such as raising their hands in response to the teacher (IASTV2--Overt Activity), doing assignments (FAIR--Solitary Work), seeking direction (IASTV2--Closed Procedural Questions), and attempting to give the correct answer to teacher questions (COS--Convergent Evaluative Interaction). In the PTE classroom, we find a more energetic, two-sided interaction pattern. Pupils demonstrate greater control both in the form of open-ended questions (IASTV2--Substantive, Open Questions) and open-ended statements (IASTV2--Open Statements and FAIR--Brings Out). Here the teacher is more interested in initiating problem solving (COS--Teacher Initiated Problem Solving) than in obtaining the routine answers to questions. Finally,

a wider range of behaviors characterizes the PTE classroom (IASTV2--Flexibility Ratio). These results are quite important in that they indicate that, consonant with its objectives, the PTE Program did produce a more pupil-centered style of interaction.

This concludes the cell by cell discussion of the significant main effects. Significant program main effects for the Student Evaluation of Teacher Training Program instrument have not been included in this discussion, since this program-evaluative instrument does not readily fit the Type of Measurement by Domain of Competence Matrix. We now turn our attention to the significant results for this instrument.

Student Evaluation of Teacher Training Program. Eight of ten scales for this instrument produced significant differences between the PTE group and the non-PTE group. Three of these main effects will not be discussed because the criterion variables in question were also associated with significant trait-treatment interactions. The remaining main effects indicate that PTE student teachers, as compared to non-PTE student teachers, found their program to be better Integrated ($p = .02$), to be a better model of Individualized Teaching ($p = .002$), to allow greater student involvement or Negotiation ($p = .06$), to provide better Teacher Preparation ($p = .02$), and to allow a greater amount of Teacher Educator--Student Teacher Interaction ($p = .02$). Recall that the scales for the Student Evaluation of Teacher Training Program instrument indicate the extent to which a teacher training program fulfills ten goals or objectives set for the PTE Program. Thus, the present results indicate that, from the viewpoint of the student teacher, the PTE Program provided an alternative to conventional teacher education programs and, furthermore, that the PTE Program seemed to be progressing toward its goals or objectives. That is, the implementation of the PTE Program was successful.

Summary of main effect results. Do the main effect results indicate general differences between the PTE and non-PTE programs? In order to answer this question, consider the overview of the main effect results given in Table 6-2. Self-report measures did provide some evidence for program differences, but this evidence tends to be limited to the interpersonal domain. For self-report measures, we find the PTE student teachers reporting themselves to be warmer, more considerate of their pupils, less worried about themselves or their pupils, and more worried about professional concerns. For other-report measures, there was no evidence for program differences, whether these other-report measures involved ratings of the trainees by their college supervisors, public school supervising teachers, or their pupils. For observational measures, there was strong evidence for program differences with respect to career-related behaviors. PTE student teachers began seeking jobs earlier, indicating greater enthusiasm concerning a teaching career, and they demonstrated a more pupil-centered style of teaching. Finally, the results for the self-report instrument entitled Student Evaluation of Teacher Training Program indicate that, according to the perceptions of the trainees, the PTE Program was successfully implemented and was progressing in the direction of its objectives.

In conclusion, the program differences for observed teaching behavior and for the student teachers' perception of their training program are quite encouraging, both indicating that the PTE Program was producing behavior change in the desired direction. Less encouraging is the lack of generality of the significant main effects. Little or no evidence of program differences was obtained for six of the eight matrix cells sampled. However, this lack of significant main effects for a

Table 6-2

Overview of Main Effect Results
in Terms of the Type of Measurement
by Domain of Competence Matrix

Type of Measurement	Domain of Competence: Intrapersonal	Interpersonal	Career-Related
Self-Report	1/11	3/11	1/11
Other-Report	0/5	0/6	0/4
Observation		0/5	10/57

Student Evaluation of Teacher Training Program = 5/10

Note: The number before the slash is the number of significant main effects; the number following the slash, the number of analyses performed.

variety of behaviors and measurement methods does not establish that there were no differential program effects of any kind for those behaviors and methods of measurement. Rather, this lack of significant main effects indicates that, for the behaviors in question, the PTE group taken as a whole was not significantly different from the non-PTE group taken as a whole. The possibility still exists that the hypothesized behavioral change may have occurred for some types of individuals even if it did not occur across all individuals. This latter contention represents the actual state of affairs, being strongly supported by a general and consistent pattern of significant trait-treatment interactions. We now turn our attention to the trait-treatment interaction results.

Chapter 7

INTERACTIONS BETWEEN PERSONALITY AND ATTITUDE TRAITS AND THE TRAINING PROGRAMS

A homogeneity of group regressions test was used to identify significant trait-treatment interactions. Recall that a significant trait-treatment interaction occurs when the relationship between the criterion variable and the trait (or predictor) variable is different for the PTE group than for the non-PTE group. Of a total of 2127 trait-treatment interaction analyses involving 111 different criterion variables, 269 analyses demonstrated significant interactions (.10 level of significance). The chance probability for obtaining at least 269 significant results is exceedingly small ($z = 4.07$; $p = .00003$), and the total set of significant trait-treatment interactions was identified as not being chance-produced.

Further attempts to identify chance-produced significant results involved consideration of the results, criterion instrument by criterion instrument. The number of significant results obtained for an instrument was compared to the number of significant results expected by chance for that instrument. If the obtained number of significant trait-treatment interactions did not significantly exceed (at the .10 level) the expected number, then significant results were identified as chance-produced. If the obtained number of significant interactions significantly exceeded (at the .10 level) the expected number, then significant results were retained as valid. (The details for this second-order significance testing are given in the preceding "Data Analysis Strategy" section of

the present chapter.) The results of these attempts to identify chance-produced significant interactions are presented in Table 7-1. A total of 134 significant interactions were retained as valid while 135 significant interactions were identified as chance-produced.

Note in Table 7-1 that the IASTV2 variables as well as the FAIR variables were divided into two separate sets for the purposes of chance-identification. This partitioning of variables was made entirely on a priori bases. For both of these instruments, one set of variables involved relatively rare behaviors, behaviors occurring less than 1% of the time, while the other set of variables involved behaviors which occurred at least 1% of the time. The variables from these two videotape coding systems were partitioned because it was felt that it would be more advantageous to apply the chance-identification procedures to smaller sets of variables. Note that the IASTV2 and FAIR were 2 of 13 criterion instruments involved in trait-treatment interaction analyses, but these two instruments provided 44% of the criterion variables for those analyses. If the number of variables in a set is large, then we run the risk of either having the effects of several potent variables masked by a large number of impotent variables or having a few potent variables carry along a large number of impotent variables. Given such impetus, a partitioning method was sought. While a random method of dividing the variables into smaller sets could have been used, a more rational plan was adopted. Clearly, a partitioning, based upon potency (i.e., a variable's ability to demonstrate the effects of program) should enhance the effectiveness of the chance-identification procedures. Finally, it was hypothesized that partitioning on the basis of different frequencies of occurrence might separate the IASTV2 and the FAIR variables with regard to potency.

Table 7-1

Number of Obtained Significant Trait-Treatment Interactions
and Chance Probability for this Obtained Number, for each Criterion Instrument

Instrument	Total Number of Trait-Treatment Interaction Analyses	Number of Results Significant at .10 Level	Number of Significant Results Expected by Chance	z	Probability	Results are Chance-Produced
Profile of Learning Priorities—Instructional Domain (Student Teacher)	88	14	8.8	1.67	.047	No
IASTV2--20 Variables[a]	440	41	44.0	<0.00	>.500	Yes
IASTV2--8 Variables[b]	176	14	17.6	<0.00	>.500	Yes
Readiness Assessment (Student Teacher)	132	14	13.2	.23	.409	Yes
Readiness Assessment (College Supervisor)	132	35	13.2	6.32	<.001	No
Readiness Assessment (Public School Supervising Teacher)	132	16	13.2	.81	.209	Yes
COS	264	28	26.4	.34	.367	Yes

Note.--Footnotes appear on the second page of this table.

Table 7-1 continued

Number of Obtained Significant Trait-Treatment Interactions
and Chance Probability for this Obtained Number, for each Criterion Instrument

Instrument	Total Number of Trait-Treatment Interaction Analyses	Number of Results Significant at .10 Level	Number of Significant Results Expected by Chance	z	Probability	Results are Chance-Produced
SET-2	66	19	6.6	4.88	< .001	No
Student Evaluation of Teacher Training Program	220	30	22.0	1.80	.079	No
Adjective Self Description[c]	7	0	.7	0.00	7.500	--
Self-Report Inventory[c]	8	0	.8	< 0.00	> .500	--
FAIR--10 Variables[d]	220	22	22.0	< 0.00	> .500	Yes
FAIR--11 Variables[e]	242	36	24.2	2.54	.006	No

[a]These are the 20 IASTV2 variables with frequency of occurrence greater than or equal to 1% of the time.
[b]These are the 8 IASTV2 variables with frequency of occurrence less than 1% of the time.
[c]When an Adjective Self Description or Self-Report Inventory scale was used as the criterion variable (post data), then the only trait-treatment interaction examined was the interaction involving the same scale (pre data) as the trait variable.
[d]These are the 10 FAIR variables with frequency of occurrence greater than or equal to 1% of the time.
[e]These are the 11 FAIR variables with frequency of occurrence less than 1% of the time.
Note.--The rationale for separating IASTV2 and FAIR variables into two separate groups is provided in the text.

The 134 significant trait-treatment interaction results retained as valid were subjected to further attempts to identify chance-produced significance. (The details for these further chance-identification procedures are presented in the preceding "Data Analysis Strategy" section of the present chapter.) These final chance-identification procedures involved two steps. First, the remaining significant results were considered, criterion variable by criterion variable. If the obtained number of significant interactions for a criterion variable did not exceed the number expected on the basis of chance, then the significant interactions for that criterion variable were discarded as chance-produced. Fifteen significant trait-treatment interaction results were discarded on this basis. Second, the 119 significant results still retained were considered, trait variable by trait variable. If the number of significant results associated with a trait variable was not greater than two, then any significant results associated with that trait variable were discarded. Twelve significant trait-treatment interaction results were discarded on this basis. Six trait variables were found to have two or fewer significant results, all of these trait variables being associated with only two significant results. The six trait variables in question were as follows:

Variable	Instrument
Anxiety	Adjective Self Description
Idealism	Adjective Self Description
Parents	Self-Report Inventory
Repetitions	One Word Sentence Completion
Hostility	One Word Sentence Completion
Anxiety	One Word Sentence Completion

Since all significant trait-treatment interaction results associated with these six trait variables have been discarded, these six trait variables will receive no further consideration in the present report.

Following all attempts to identify chance-produced results, 107 trait-treatment interaction results were retained as valid. The chance-identification procedures were quite conservative, involving successive tests on four levels--1) the total set of trait-treatment interaction analyses, 2) each criterion instrument, 3) each criterion variable, and 4) each trait variable. The conservative nature of the chance-identification procedures allows a high level of confidence to be placed in the remaining significant results. The remainder of the present chapter will be devoted to a discussion of these remaining significant trait-treatment interaction results.

Appendix B presents the 107 significant trait-treatment interaction results retained as valid. Our real interest is not in merely cataloguing the significant results but rather in identifying any general patterns or trends associated with them. To these ends, the following discussion has been divided into three sections. The first section provides a framework for discussing the consistency of trait-treatment interaction results. The second section deals with the generality of the obtained trait-treatment interactions with regard to the criterion variables involved. The distribution of the significant results within the Type of Measurement X Domain of Competence Matrix will be examined in order to determine if the significant findings tend to be concentrated with regard to specific types of measurement and/or specific kinds of behavior. In addition, the consistency of the form of the significant interactions will be considered for each relevant criterion variable, for each matrix cell, and overall.

The third section deals with the generality of the obtained trait-treatment interaction results with regard to the trait variables involved. This final section, entitled "Predictive Efficiency," will focus on the extent to which each trait variable defines regions of significant differences between programs.

A Framework for Discussing the Consistency of Trait-Treatment Interaction Results

In order to deal with the question of the consistency of the form of the obtained interactions, it is necessary to define general forms that the interactions may assume. This dictates that a quite lengthy set of introductory remarks precede the actual description of results. While the framework provided in the next few pages may seem somewhat complicated, such a framework proves its worth in the discussion of the trait-treatment interaction results.

Types of trait-treatment interactions.[1] Four general types of trait-treatment interaction will be considered. A type Pn interaction is one for which the trait-criterion relationship is relatively more positive for the PTE Program than the non-PTE. That is, the slope (beta-weight) for

[1] This discussion of types of trait-treatment interactions is limited to interactions that are disordinal. A trait-treatment interaction is disordinal if the trait-criterion regression lines for the two programs intersect within the range of the data, with a resulting superiority on the criterion measure for one program for high trait values but a superiority for the other program for low trait values. On the other hand, an ordinal interaction occurs when the regression lines for the two programs do not intersect within the range of the data, one program always being superior. Limiting consideration to disordinal interactions simplifies the discussion of types of interaction and presents no problems in the case of the present data since 106 of the 107 significant trait-treatment interactions were disordinal.

the regression line of the criterion variable on the trait variable is greater (more positive) for the PTE Program than for the non-PTE program, the PTE Program evidencing superior criterion performance for high trait values but the non-PTE program being superior for low trait scores. A type P+/n- interaction is an interaction for which the trait-criterion relationship is positive for the PTE Program and negative for the non-PTE program. Note that all P+/n- type interactions are also Pn type interactions, but the reverse is not true. That is, P+/n- interactions form a subset of Pn interactions.

A type pN interaction is one for which the trait-criterion relationship is relatively more positive for the non-PTE program than for the PTE Program. That is, the slope for the regression line of the criterion variable on the trait variable is greater (more positive) for the non-PTE program than the PTE Program, the non-PTE program demonstrating superior criterion performance for high trait scores but the PTE Program being superior for low trait scores. A type p-/N+ interaction indicates that the trait-criterion relationship is positive for the non-PTE program but negative for the PTE Program. Note that p-/N+ interactions form a subset of the pN interactions.

The examples given in Figure 7-1 should clarify the meanings of the four types of interactions. In order for these four types of interactions to be meaningful, it is necessary that all trait variables be scaled in a consistent manner and that all criterion variables be scaled in a consistent manner. Consider the situation when scaling is not consistent. Say that one criterion variable is a negatively scaled variable such as SET 2--Rapport (a high score representing low teacher-pupil rapport and a low score representing high rapport), while another criterion

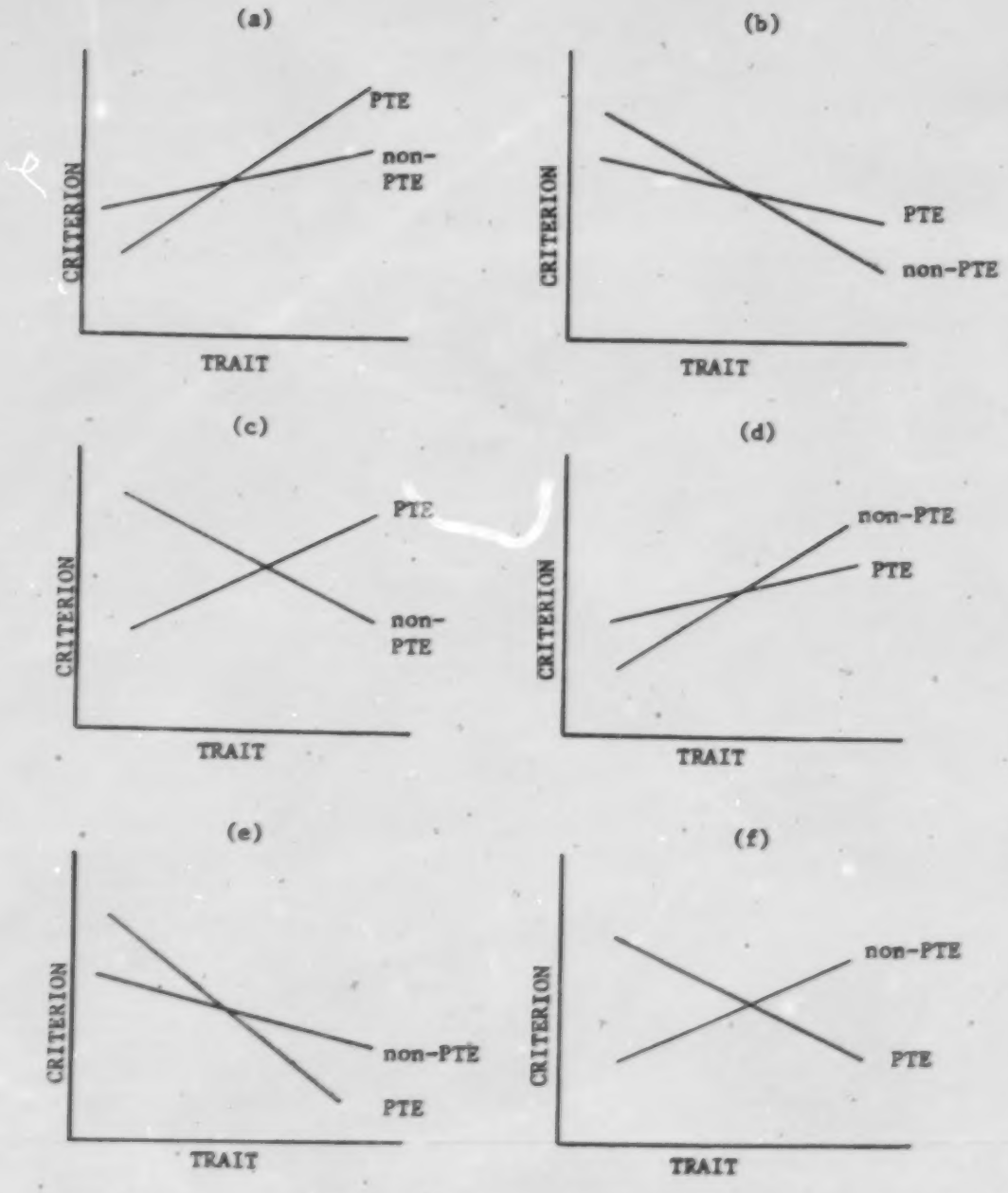


FIG. 7-1. Different types of trait-treatment interactions. (a) a Pn type interaction with the trait-criterion relationships positive for both programs. (b) a Pn type interaction with the trait-criterion relationships negative for both programs. (c) a Pn type interaction which is also P+/n-. (d) An nP type interaction with the trait-criterion relationships positive for both programs. (e) An nP type interaction with the trait-criterion relationships negative for both programs. (f) an nP type interaction which is also p-/N+.

variable is a positively scaled variable such as FAIR--Student Explores (a high score representing more exploration than a low score). The meanings of the four types of interactions will be different for these two criterion variables. Inconsistency of the scaling of the trait variables will produce analogous problems.

Criterion variable scaling. Only those criterion variables for which significant trait-treatment interaction results were obtained are considered with regard to scaling consistency. The following method was used to ensure comparability of scaling for the criterion variables. 1) Consider criterion variables representing positive behavior. If the scaling was such that a high score represented less intensity or a lower frequency of occurrence for the behavior (and a low score represented greater intensity or a higher frequency of occurrence for that behavior), then the scaling was reversed. Otherwise, the scaling was not altered. 2) Consider criterion variables representing negative behavior. If the scaling was such that a low score represented less intensity or a lower frequency of occurrence for the behavior (and a high score indicated greater intensity or a higher frequency of occurrence for that behavior), then the scaling was reversed. Otherwise, the scaling was not altered.

These methods led to the reversal of the scaling for four of the 22 criterion variables for which significant trait-treatment interactions were found. The scaling for two SET 2 variables (Rapport and Fosterance of Self-Esteem) was reversed, since a high score on these two variables indicated a low intensity of the behaviors in question. The scaling of two variables (Teacher Is Tangential and Teacher Solitary Work) from the FAIR videotape scoring system was also reversed. The behaviors classified

by these two variables were deemed to be negative but the scaling in both cases was appropriate to positive behaviors.

Following the reversal of scaling for these four criterion variables, the scaling of all 22 relevant criterion variables was such that a high score corresponded to a more positive behavior, and a low score corresponded to a more negative behavior. Note that the decision to rescale the other 18 relevant criterion variables was made prior to the inspection of the types of interaction found for each variable.

Trait variable scaling. An a priori decision was made to scale the trait variables on the basis of their desirability. If a trait variable was thought to represent a characteristic which is desirable, then the scaling for that variable was not altered. If a trait varaible was thought to represent a characteristic which is undesirable, then the scaling for that variable was reversed. Finally, if the relevance of a trait variable was unclear, then that trait variable was eliminated from discussions dealing with the types of obtained trait-treatment interactions. Table 7-2 presents the classification of the trait variables with regard to their relevance.

It was necessary to determine relevance only for the 16 trait variables for which non-chance significant trait-treatment interaction results were found, and these 16 trait variables are those presented in Table 7-2. The scaling for the four variables classified as undesirable was reversed. Following this reversal, the 14 relevant traits (both desirable and undesirable) were all scaled such that a high score indicated greater desirability than a low score.

General comments. Note that the significant trait-treatment interactions listed and described in Appendix B are based upon the original

Table 7-2

Classification of the Trait Variables
With Regard to their Relevance

Desirable Traits	Desirable Traits	Desirable Traits
Adjective Self Description:	Adjective Self Description:	One Word Sentence Completion:
Attitude	Behavior	Populars
Efficiency	Introversion	Response Length
Attractiveness	One Word Sentence Completion:	
Self-Report Inventory:	Evasion	
Self	Depression	
Others		
Children		
Work		
Reality		
Hope		
Authority		

scalings for all variables and not upon any reversed scalings. For the present purposes, the only effect of reversing the scaling of a variable was to change the signs of the beta-weights and correlation coefficients associated with that variable.

The four types of trait-treatment interaction defined in the present section should provide a framework for discussing the consistency of the trait-treatment interaction results. The results are consistent to the extent that the trait-treatment interactions which occur for the different criterion variables tend to be of the same type. We now turn our attention to the consistency and generality of the trait-treatment interaction results across different criterion variables.

Consistency and Generality of the Trait-Treatment Interaction Results Across Criteria

The significant trait-treatment interaction results will be discussed below according to their classification in terms of the Type of Measurement X Domain of Competence Matrix. Each of the nine cells will be considered in turn. In the following discussion, interactions involving the 14 relevant trait variables have been classified according to type. Interactions involving the irrelevant traits (Populars and Response Length from the One Word Sentence Completion) cannot be classified as to type and are discussed separately.

Cell 1: Self-Report X Intrapersonal Behavior. Of the 51 trait-treatment interaction analyses performed for this cell, no analysis detected a non-chance significant interaction.

Cell 2: Self-Report X Interpersonal Behavior. Fifty-one trait-treatment interaction analyses were also performed for this cell. Again there was no evidence for a non-chance significant interaction.

Cell 3: Self-Report X Career-Related Behavior. Of the 133 analyses performed for this cell, 13 provided evidence of non-chance significant interactions. All 13 of these significant results were associated with the variables from the Profile of Learning Priorities instrument. For the Competent Management criterion variable, a significant interaction was found between program and the Populars scale from the One Word Sentence Completion--the form of this interaction being a negative trait-criterion relationship for the PTE Program but a positive relationship for the non-PTE program.

The remaining 12 significant interactions have been classified as to type in Table 7-3. Of these 12 interactions, 8 are of the pN variety, while only 4 are of the Pn variety. There is a moderately strong tendency for the interactions in cell 3 to be of the pN variety. That is, the relationship between desirable characteristics, on the one hand, and self-reported flexibility, professionalism, responsibility, and competent management, on the other hand, tends to be more positive for non-PTE trainees than for PTE trainees.

Cell 4: Other-Report X Intrapersonal Behavior. Of a total of 110 analyses for this cell, 17 were found to produce evidence for significant trait-treatment interactions. The program by Response Length interaction was significant for both the Perceptive About Self and the Self-Concern variables from the Readiness Assessment (as completed by the College Supervisors). Both interactions were produced by a positive trait-criterion relationship for the PTE Program but a negative trait-criterion relationship for the non-PTE program.

Table 7-3

Significant Trait-Treatment Interactions
for Criterion Variables in Cell 3--
Self-Report Measures by Career-Related Behaviors

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
Profile of Learning Priorities					
1) Flexibility	2	1	1	1	1
2) Professionalism	3	1	1	2	0
3) Responsibility	3	2	2	1	1
4) Competent Management	4	0	0	4	2
Total	12	4	4	8	4

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion-trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

The remaining 15 significant interactions have been classified as to type in Table 7-4. All of these 15 interactions are of the pN type, and 14 of the 15 are of the p-/N+ type. That is, college supervisors (Readiness Assessment) and pupils (SET II) evaluated the personal characteristics of PTE trainees more negatively if those trainees scored high on desirable traits than if those trainees scored low on desirable traits. The opposite relationship was true for the non-PTE student teachers, high scores on desirable traits leading to higher evaluations from pupils and college supervisors.

Cell 5: Other-Report X Interpersonal Behavior. Twenty-six of the 132 analyses for this cell produced evidence of significant trait-treatment interactions. The program by Response Length interaction was found to be significant for both Fosterance of Self-Esteem (SET II) and Perceptive About Children (Readiness Assessment as completed by the College Supervisor). In both cases, this interaction is produced by a positive trait-criterion relationship for the PTE Program but a negative relationship for the non-PTE program. The program by Populars interaction was found to be significant for the Fosterance of Self-Esteem variable as well as the Concern for Children variable (Readiness Assessment--College Supervisor). For both crieria, this interaction was produced by a negative trait-criterion relationship for the PTE Program and a positive relationship for the non-PTE program.

The remaining 22 interactions have been classified as to type in Table 7-5. All but one of these significant interactions are of the pN type, and 20 of these 21 are of the p-/N+ type. Note that this is the same consistent pattern established for cells 3 and 4. For cell 5, college supervisors (Readiness Assessment) and pupils (SET II) evaluated the inter-

Table 7-4

Significant Trait-Treatment Interactions
for Criterion Variables in Cell 4--
Other-Report Measures by Intrapersonal Behaviors

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
Set 2--Unreasonable Negativity	3	0	0	3	3
Readiness Assessment (College Supervisor)					
1) Perceptive about Self	5	0	0	5	5
2) Self Concern	7	0	0	7	6
Total	15	0	0	15	14

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

Table 7-5

Significant Trait-Treatment Interactions
for Criterion Variables in Cell 5--
Other-Report Measures by Interpersonal Behaviors

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
Set 2					
1) Rapport	9	0	0	9	9
2) Fosterance of Self-esteem	4	1	1	3	3
Readiness Assessment (College Supervisor)					
1) Concern for Children	4	0	0	4	4
2) Perceptive about Children	5	0	0	5	4
Total	22	1	1	21	20

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

personal skills of PTE trainees more negatively if those trainees scored high on socially desirable traits than if those trainees scored low on these traits. The opposite relationship was true for the non-PTE student teachers, high scores on desirable traits leading to higher evaluations from pupils and college supervisors.

Cell 6: Other-Report X Career-Related Behavior. Nine of 88 analyses produced significant results for this cell. The program by Response Length interaction was significant for the Concern for Impact criterion (Readiness Assessment--College Supervisor), this interaction being produced by a positive trait-criterion relationship for the PTE Program but a negative relationship for the non-PTE program.

The eight remaining significant interactions are classified as to type in Table 7-6. All of these 8 interactions are of the pN variety, and 7 of these are of the p-/N+ type. Thus, the same consistent pattern is repeated in cell 6. In cell 6 we find college supervisors perceiving greater Concern for Impact for PTE students who score low on desirable traits than for PTE students who score high on these traits. On the other hand, non-PTE trainees received higher Concern for Impact ratings if they scored high rather than low on desirable traits.

Cell 7: Observation X Intrapersonal Behavior. The present investigation did not attempt to study any variables which would be classified within this category. Only physiological measures would provide criterion variables for this cell.

Cell 8: Observation X Interpersonal Behavior. No significant results were found for the 110 analyses performed for this cell.

Table 7-6

Significant Trait-Treatment Interactions
for Criterion Variables in Cell 6--
Other-Report Measures by Career-Related Behaviors

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
Readiness Assessment (College Supervisor)-- Concern for Impact	8	0	0	8	7
Total	8	0	0	8	7

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

Cell 9: Observation X Career-Related Behavior. For this cell, 23 of the 1232 analyses produced significant results. While significant results were obtained for this cell, the results are not overly encouraging since only 1.9% of the analyses produced significant results. The program by Populars interaction was significant twice--for the Teacher is Tangential and the Student Explores variables from the FAIR coding system. One of these interactions was produced by a positive relationship for PTE and a negative relationship for non-PTE, but the other was produced by the reverse relationships. The program by Response Length interaction was significant for the Student Admits (FAIR) criterion variable, this interaction being caused by a positive trait-criterion relationship for the PTE Program but a negative relationship for the non-PTE program.

All of the remaining 20 significant interactions were associated with criterion variables from the FAIR instrument. These remaining significant interactions are classified as to type in Table 7-7. Eighteen of these 20 interactions are of the pN form, with 16 of these 18 also being p-/N+. Although the number of significant results obtained for this cell was not very encouraging, the significant results that were found provided the same consistent pattern obtained in previously discussed cells. The significant results obtained for the present cell indicate that a lower frequency of Teacher Solitary Work and Tangential behavior by the teacher and a higher frequency of Student Admits, Rejoices for Self, Questions and Explores occurred for PTE trainees who scored low on desirable traits than for PTE trainees who scored high on these traits. On the other hand, the classrooms of non-PTE trainees evidenced a lower frequency of these teacher behaviors and a

Table 7-7

Significant Trait-Treatment Interactions
for Criterion Variables in Cell 9--
Observational Measures by Career-Related Behaviors

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
PAIR					
1) Teacher Solitary Work	4	0	0	4	4
2) Teacher is Tangential	5	0	0	5	5
3) Student Admits	3	0	0	3	2
4) Student Rejoices for Self	3	0	0	3	2
5) Student Questions	3	1	1	2	2
6) Student Explores	2	1	1	1	1
Total	20	2	2	18	16

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

higher frequency of these student behaviors if the trainees scored high rather than low on desirable traits.

This concludes the cell by cell discussion of the significant trait-treatment interaction results. Significant trait-treatment interactions for the Student Evaluation of Teacher Training Program instrument have not been included in this discussion, since this program-evaluative instrument does not readily fit the Type of Measurement X Domain of Competence Matrix. We now turn our attention to the significant results for this instrument.

Student Evaluation of Teacher Training Program. For this instrument, 19 of the 220 analyses produced significant results. All 19 of these significant interactions involved behaviorally relevant trait variables, and these significant interactions have been classified as to type in Table 7-8. Five of the interactions are of the Pn type. The 14 pN type interactions include 13 which are also p-/N+. Overall, the results for the Student Evaluation of Teacher Training Program instrument indicate that PTE trainees rated their program higher if those trainees scored low on desirable traits, while non-PTE trainees rated their program higher if those trainees scored high on desirable traits.

Summary. The consistency of the obtained significant trait-treatment interactions is nothing short of amazing. (An overview of the significant trait-treatment interactions is presented in Table 7-9.) Of the total of 96 interactions classifiable as to type, 84 were of the pN variety, and 74 of these 84 were also p-/N+. For the PTE Program, the individuals who performed the best and received the highest evaluations were those who lacked desirable traits. For the non-PTE program, the

Table 7-8

Significant Trait-Treatment Interactions
for Criterion Variables from the
Student Evaluation of Teacher Training Program Instrument

Criterion Variables	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
1) Person-Centered	3	0	0	3	3
2) Personal, Intellectual and Social Development	4	4	3	0	0
3) Behavior Modelling	12	1	1	11	10
Total	19	5	4	14	13

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

Table 7-9

Overview of Types of Significant Trait-Treatment Interactions in Terms of the Type of Measurement By Domain of Competence Matrix

Matrix Cells	Total # of Significant Trait-Treatment Interactions	Types of Trait-Treatment Interactions			
		Pn	P+/n-	pN	p-/N+
1) Self-Report x Intrapersonal	0	0	0	0	0
2) Self-Report x Interpersonal	0	0	0	0	0
3) Self-Report x Career-Related	12	4	4	8	4
4) Other-Report x Intrapersonal	15	0	0	15	14
5) Other-Report x Interpersonal	22	1	1	21	20
6) Other-Report x Career-Related	8	0	0	8	7
7) Observation x Intrapersonal	--	--	--	--	--
8) Observation x Interpersonal	0	0	0	0	0
9) Observation x Career-Related	20	2	2	18	16
Subtotal	77	7	7	70	61
Student Evaluation of Teacher Training Program	19	5	4	14	13
Total	96	12	11	84	74

Explanation of trait-treatment interaction types:

Pn. These are interactions where the criterion trait regression is relatively more positive for the PTE Program than for the Non-PTE Program.

P+/n-. This is a subset of the Pn interactions where the criterion is positively related to the trait for the PTE Program and negatively related to the trait for the Non-PTE Program.

pN. The reverse of the Pn interaction.

p-/N+. The reverse of the P+/n- interaction.

individuals who performed the best and received the highest evaluations were those who already possessed desirable traits.

The 12 Pn type interactions obtained were examined to determine if these interactions were produced by a few specific trait variables. However, no trait variable was involved in more than three of these interactions, and these interactions were divided among seven of the fourteen relevant trait variables.

An examination of the significant results for the two behaviorally irrelevant trait variables indicates that 4 of 5 results for the program by Populars interaction were produced by positive trait-criterion relationships for the non-PTE program and negative relationships for the PTE Program, while 4 out of 4 results for the program by Response Length interaction were caused by positive trait-criterion relationships for the PTE Program but negative relationships for the non-PTE program. If the scaling for the Response Length variable were reversed and the interactions for the two behaviorally irrelevant trait variables then classified as to type, then 8 of the 9 obtained significant interactions would be of the p-/N+ form. Thus, the original scaling for the Populars variable and the reversed scaling for the Response Length variable produced the same consistent pattern of results as found for the behaviorally relevant trait variables. This indicates that a high score on the Populars variable but a low score on the Response Length variable (original scaling) may be behaviorally desirable.

While the consistency of the trait-treatment interaction results is quite encouraging, the generality of these results is not as impressive. Table 7-10 presents the cell by cell percentage of results accepted as significant. Inspection of this table indicates that the trait-treatment

interaction results are strongest for other-report measures. Outside of the Other-Report row of the matrix, the only strong results are for the Self-Report X Career-Related cell. While the generality of the trait-treatment interaction results is somewhat limited, it is quite interesting that these results serve to complement the main effect results. Recall that the main effect results were strongest for the Self-Report X Interpersonal cell and the Observation X Career-Related cell. (See Table 6-2.) These are two cells where the trait-treatment interaction results are especially weak. Similarly, there were no significant main effects for any other-report measures, but these measures provide the strongest trait-treatment interaction results.

Considering the main effect and trait-treatment interaction results together, there is quite general evidence for the differential influence of the two programs. Strong evidence for such influence is provided except in the cases of the Self-Report X Intrapersonal cell and the Observation X Interpersonal cell.

Predictive Efficiency

This final section of the present chapter deals with two basic issues. First, to what extent do the trait-treatment interaction results allow prediction of program differences for different types of individuals? Do regions of significant differences between the two programs exist for the significant trait-treatment interactions? Here interest is focused upon the results of the Johnson-Neyman analyses for regions of significance. Second, what is the predictive efficiency of each of the trait variables? Are some trait variables better predictors of program differences than others?

Table 7-10

Percentage of Trait-Treatment Interaction Results
Accepted as Significant in Terms of the Type of Measurement
By Domain of Competence Matrix

Type of Measurement	Domain of Competence: Intrapersonal	Interpersonal	Career-Related
Self-Report	0.0%	0.0%	9.8%
Other-Report	15.4%	19.7%	10.2%
Observation		0.0%	1.9%

Student Evaluation of Teacher Training Program = 8.6%

Note: The tabled percentages = $\frac{\text{\# of significant results for a cell}}{\text{total \# of analyses for that cell}} \times 100$

Results of the Johnson-Neyman analyses. The results of the Johnson-Neyman analyses (Appendix B) indicate that regions of significance existed for 67% of the significant interactions. Lower regions of significance were found for 55 of the 107 significant interactions, while 34 upper regions of significance were found. Note that regions of significance were defined within the range of the existing data. In other words, each region of significance included at least 1 subject. Coupled with the fact that most interactions were of the p-/N+ type, finding this many regions of significance implies that, for individuals who scored low on desirable traits, the PTE training program was superior to the non-PTE training program. On the other hand, there was some tendency for individuals who scored high on desirable traits to profit more from the non-PTE training program than the PTE trianing program.

Table 7-11 presents the significant trait-treatment interaction results classified as to trait variable. This table presents the number of significant results for each trait variable, the number of results for which regions of significance occurred, and information relevant to the characteristics of the obtained regions of significance. An "x" appears in the "a" column of this table if lower regions of significance predominate, and in the "b" column if upper regions predominate, and in the "c" column if no consistent pattern emerges.

Seven of the trait variables demonstrated a predominance of lower regions while only four demonstrated a predominance of upper regions. Similarly, the median percentage of cases in the lower region exceeded that in the upper region for 11 of the 16 trait variables. Program differences for low trait values were stronger than program differences for

Table 7-11

Predictor Efficiency

Predictor Variable	No. Criteria for which TTI Present	No. Criteria for which there are Regions of Significance	a[1]	b[2]	c[3]	Median % Cases (Ss) in Lower Region[4]	Median % Cases (Ss) in Upper Region[4]
ASD/Attitude	11	8	x			13	0
ASD/Behavior	6	4		x		0.5	30
ASD/Efficiency	8	5			x	12	9
ASD/Introversion	7	4	x			25	2.5
ASD/Attractiveness	6	5			x	21	3
SRI/Self	4	3	x			9	1
SRI/Others	10	7	x			10	0
SRI/Children	7	5	x			18	0
SRI/Authority	5	1		x		3	20
SRI/Work	11	10			x	16.5	0

Note.--Footnotes appear on the second page of this table.

Table 7-11 continued

Predictor Efficiency

Predictor Variable	No. Criteria for which TTI Present	No. Criteria for which there are Regions of Significance	a[1]	b[2]	c[3]	Median % Cases (Ss) in Lower Region[4]	Median % Cases (Ss) in Upper Region[4]
SRI/Reality	4	3			x	0	9
SRI/Hope	4	2	x			7.5	0
OWSC/Response Length	6	2		x		0	32
OWSC/Populars	5	4	x			27.5	0
OWSC/Evasion	6	4		x		0	18
OWSC/Depression	5	3			x	8	2

[1]An x appears in this column if most of the program differences tend to occur at the lower region of significance.
Note: This indicates that for 66% or more of the criteria for which the predictor has regions of significance, 75% or more of the cases in those regions occur in the lower region.
[2]An x appears in this column if most of the program differences tend to occur at the upper region of significance.
Note: This indicates that for 66% or more of the criteria for which the predictor has regions of significance, 75% or more of the cases in those regions occur in the upper region.
[3]An x appears in this column if there is no consistent trend to the differences.
[4]Median % of cases refers to the median % for those criteria for which regions of significance exist.

high trait values. Recall that the PTE program tended to be superior for low trait scores, but the non-PTE program tended to be superior for high trait scores. Thus, the superiority of the PTE program for low trait values was somewhat more substantial than the non-PTE superiority for high trait values.

Efficiency of the trait variables as predictors. To the extent that certain trait variables are effective predictors of program differences, these trait variables can be used to assign individuals to the program from which they should profit most. Information relevant to the effectiveness of the trait variables as predictors is presented in Table 7-11. Three of the trait variables studied in the present investigation stand out with regard to the number of criteria for which predictions of program differences is possible. The Attitude variable was involved in a significant interaction for 11 of the criterion variables and regions of significance were obtained for 8 of these 11 criteria. Eight lower regions of significance and two upper regions of significance were found. The use of the Attitude variable for assignment of individuals to programs would thus be mainly limited to assigning students who score low on this trait to the PTE program. The Others variable was involved in 10 significant interactions and regions of significance were found for 7 of these interactions. Seven lower regions of significance and one upper region occurred for this variable. The predictive usefulness of the Others variable is also limited to assigning students with a low Others score to the PTE Program. The Work variable was involved in 11 significant interactions with regions of significance occurring for 10 of these interactions. Ten lower regions of significance and four upper

regions were obtained. Thus, this variable shows promise for assignment of low scoring individuals to the PTE Program and, to a lesser extent, the assignment of high scoring individuals to the non-PTE program.

This concludes the presentation of the results of the present investigation, The following chapter traces the conceptual development of Personalized Teacher Education and concludes this report by placing the results of this study in historical context.

SECTION III.

CONCLUSION

Chapter 8

Evaluating Personalized Teacher Training: Some Concluding Remarks

This chapter represents a mix of random thoughts and systematic observations about the development and evaluation of the personalized model of teacher training. The first section of this chapter deals with the conceptual development of the Personalized Teacher Education Program and, in a retrospective manner, recounts the program's conceptual and theoretical history from roughly the period 1971-1974. A second section deals with methods for evaluating the Personalized Teacher Education Program and, in specific, presents several formative strategies for evaluating and revising the program. A third technical section has been included in order to detail several statistical problems that, while yet unresolved, have been brought to the foreground by the present study. Interwoven throughout the chapter are the authors' candid impressions of the strengths and weaknesses of the present study.

Some Comments About the Conceptual Development of the Personalized Model of Teacher Training

In the life of every program with which the authors have worked, there have been definable stages of growth that, when viewed in sequence, represent an ontological or developmental view of that program. The Personalized Teacher Education Program (PTE) is no exception as its developing framework, both practical and theoretical, can be noted in writings as early as the Mental Health and Teacher Education Study (Peck, 1958) and the Personality, Teacher Education, and Teaching Behavior Study (Fuller, Peck, et al., 1969). Many of the major conceptual elements of the PTE Program have been drawn from these initial writings and, over the course of time, modified and expanded as the

theory or concept of personalization has grown into practice. From this and later work we can identify four distinct developments in the growth of the Personalized Teacher Education Program: its development as (a) a metatheory or global framework guiding the development of sociopsychological theories of intrapersonal and interpersonal behavior, (b) a theory or set of theories with which relationships between intra- and interpersonal behaviors, the concepts of mental health and teaching effectiveness can be posited, (c) a prototypic model for teacher training and (d) a specific, i.e., ongoing, example of a personalized teacher training program. The following comments focus upon the current developmental status of Personalized Teacher Education vis-a-vis the concepts of metatheory, theory, model and example.

Metatheory. A metatheory is a grammar or medium of communication that is concerned with the development, investigation or description of a theory. It is used to specify the rules with which a theory is to be constructed and formulated, e.g., according to some particular philosophical or logical system. PTE, since its earliest stages, has embodied implicit rules for the construction of theory. Some of these have been discussed in the opening chapter of this report and include, for example, the intrapersonal, interpersonal and career-related domains as the sine qua non of personalization.

Early in its development, research on Personalized Teacher Education suggested the importance of these domains for training prospective teachers. The first domain is in many respects the most important and concerns the acquisition of intrapersonal skills. Intrapersonal skills are competencies that a prospective teacher acquires that assist her in learning about her own abilities and emotions as these relate to effective teaching. Such psychological constructs as self-confidence, self-perception and the congruence between one's own feeling and behavior are posited by the

PTE Program as important correlates of effective teaching. The second domain of competence, that of interpersonal skills, also is inextricably woven into the philosophy of Personalized Teacher Education. Early developmental work (Peck, 1958) has stressed the importance of training teachers to be able to relate to others, to be responsive, to be able to show empathy, to be receptive to feedback and to possess supportive attitudes toward their pupils. These concepts have subsequently become known as coping skills and strategies (Peck, 1971). Thirdly, the Personalized Teacher Education Program embodies what has become the mainstay of many of the so-called traditional training programs, that of career-related behaviors. Career-related behaviors emphasize the cognitive skills and knowledge most directly related to the act of teaching but may also include such related skills as classroom management, knowledge of child development, skills in pupil evaluation and alternative teaching styles.

The basic philosophy of Personalized Teacher Education has rested upon the extant relationships among the intrapersonal, interpersonal and career-related domains. This philosophy or metatheory stands in contradistinction to other metatheoretical systems that would emphasize or seek to develop one domain of competence to the exclusion of others. By positing a global framework for teacher training from these domains, Personalized Teacher Education has contributed a grammar and a medium of communication for the purpose of developing the specific theories and concepts of personalization. These basic domains of competence represent the philosophy or logic system by which the Personalized Teacher Education Program has developed.

Theory. A theory is a set of conceptual units and a schema for the interrelationship of these units. In simplest terms, a theory is a symbolic construction designed to bring generalizable facts, concepts or variables into

systematic connection. These facts, concepts or variables, then, are defined and used to make empirical and theoretical predictions about behavioral events. For example, Mandler and Kessen (1959) find that the purpose of a theory is similar to that of a road map:

> The road map is an artificial, symbolic and reduced representation (a theory) of the terrain and the schooled reader of the map may act in a reasonable way (behave functionally, behave factually) over that terrain with the help of the map. The rules for interpretation of the map correspond in a rough way to definition and theory construction.

In its purest form, a theory usually consists of hypotheses or axioms, a mathematical system for testing hypotheses, a technical vocabulary and a model or working example. While Personalized Teacher Education has not engendered any theories with all of these ingredients, it has provided the impetus for and the development of what has become known within Personalized Teacher Education and in recent literature as the concerns theory (Fuller, 1969a; Fuller, 1974).

The concerns theory conceptualizes the learning process for a prospective teacher as a natural flow from concerns for self (trainee) to task (teaching) to impact (pupil). Since the theory posits that learning in this sequence proceeds from the self, the prospective teacher is the starting point for planning and structuring learning experiences within the Personalized Teacher Education Program. While the concerns theory has a history and development of its own, the philosophy and logic system of Personalized Teacher Education, i.e., the metatheory, led to its initial conceptualization and later development. The concerns theory operationalizes the concepts of intrapersonal, interpersonal and career-related behavior for the purpose of moving the prospective teacher from a focus on self (intrapersonal) to a focus on the teaching task (interpersonal) and, ultimately, to a focus on the impact she is having upon pupils (career-related).

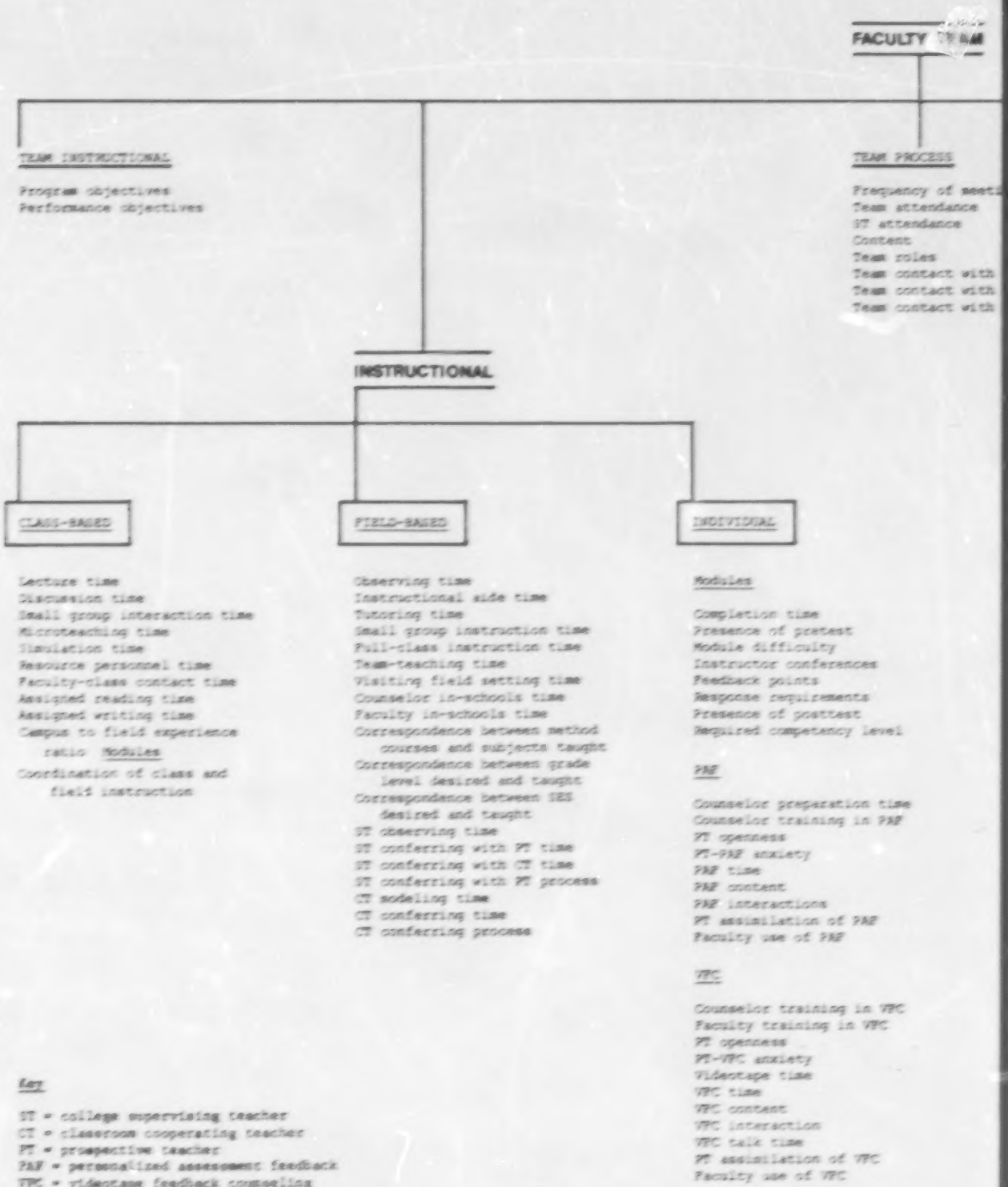
While a metatheory is typically an abstract conceptual tool used in the development of theory, theory itself consists of hypothetical, i.e., abstract concepts, observed data and intervening variables that link hypothetical concepts to observed data. Theories that deal primarily with hypothetical concepts and intervening variables are referred to as conceptual theories, of which the concerns theory is an example.

A second kind of theory that has been used in the Personalized Teacher Education Program is that of the descriptive theory. Descriptive theories have been used to graphically define variables and the relationships among them in order to reduce them to their most elementary and concrete form. The processes of classification, grouping and variable definition are used in the development of descriptive theories. Systematic descriptions of observable phenomena are constructed through the use of these processes for the purpose of devising working models and examples of the program. One such descriptive theory of Personalized Teacher Education constructed early in its development appears in Figure 8-1.

Note that the figure portrays crude relationships between the three major concepts: use of the faculty-team approach, application of a broad-based instructional program and attention to the personal needs and histories of program participants. Each of the six more elementary units of the descriptive scheme appearing below these major concepts are defined with specific behavioral variables, reducing each program component to its most elementary and concrete form. These definitions and variables qualitatively have assisted developers in relating the operational components of the PTE Program to the more general conceptual and hierarchical concepts and to the metatheory itself.

Model. A third element in the ontological development of Personalized Teacher Education has been the use of models. PTE has employed two types of

Figure 8-1. Descriptive Theory of the Intrapersonal, Interper



...personal and Career-Related Behaviors Operating in PTE

...ETI

...eetings
...e

...ith ST
...ith CT
...ith PT

TEAM PERSONAL

Teaching experience
Teacher education experience
PTEP experience
Team PTE in PTEP
Counselor's experience--general
Counselor's experience--PTs

PERSONAL

PROSPECTIVE TEACHER

Biographical

Sex
Age
Marital status
GPA
Scholastic Aptitude
Area of Concentration
Semester in college

Personality

Dominance
Sociability
Self-acceptance
Conformity
Achievement via conformance
Achievement via independence
Anxiety

Attitude

Concerns level
Attitude toward teaching
Attitude toward PTEP
Attitude toward CT
Attitude toward ST
Performance Objectives

SUPERVISING TEACHER

Biographical

Supervisory experience
Graduate credit hours
Current PT load
Teaching experience

Personality

ST-PT compatibility

Attitude

Concerns level
Attitude toward teaching
Attitude toward PTEP
Performance objectives for PT

COOPERATING TEACHER

Biographical

Secondary teaching experience
Elementary teaching experience
Experience as CT
Degree status

Personality

CT-PT compatibility

Attitude

Concerns level
Attitude toward teaching
Attitude toward PTEP
Performance objectives for PT

models, symbolic and pictorial. A symbolic model, in its purest form, describes behavioral properties in coded, usually mathematical, terms. The psychologist Kurt Lewin, for example, used symbolic models for describing behavior so that observations such as "the child chose the toy he liked best" could be expressed with such symbolic formulae as "$F_p G_1 > F_p G_2$" meaning the child liked Goal 1 (G_1) better than (>) Goal 2 (G_2) and that personality characteristics (p) of the child could be used to account for his behavior. Symbolic models in the development of the Personalized Teacher Education Program have often taken the form of what Royce (1963) and others (Margenau, 1950) have called nomological networks. Nomological networks--contrary to descriptive networks--describe the internal relationships between behaviors within a program in *hypothetical* and *schematic* form. Unfortunately, many relationships between instructional variables and criterion behaviors have yet to be empirically confirmed so that no mathematical system is appropriate of the sort that would allow us to predict that for every unit increase in a given instructional variable, we can expect x units of increase in a corresponding criterion behavior. Much more prominent in Personalized Teacher Education than symbolic models has been the idea of a pictorial model that depicts a graphic sequence of events much like a PERT chart of an instructional program portrays major milestones of program activities and accomplishments. A pictorial chart of this type was presented in the context of Chapter 2 in order to show the sequence in which prospective teachers receive components of the Personalized Teacher Education Program. Pictorial models are usually graphic, not mathematical, representations and are often characteristic of a program for which causal relationships between instructional variables and criterion behaviors have yet to be confirmed.

Examples. The fourth element which has played a role in the development of Personalized Teacher Education is that of the example. Much as models are built from theories, examples result from and are illustrations of models, either symbolic or pictorial. Examples are simply working representations of all the conceptualization that has gone before in the developmental sequence, i.e., metatheory, theory and models. Examples are most appropriate when fully developed metatheories have led to sound descriptive and conceptual theories that, in turn, have led to the empirical validation of either a symbolic or pictorial model. Examples are the last and most concrete elements in the ontological sequence and should embody the spirit of the metatheory and its theoretical concepts and models. While there can be only one metatheory undergirding the ontological sequence, we expect and hope that numerous theories and as many models and examples as are needed to describe, to test and to demonstrate the metatheory are developed. While a single metatheory holds the system together, multiple theories, models and examples work side by side to make operational the philosophy and logic system of the metatheory. If the metatheory, theory and models are well developed, the example is most likely to effectively demonstrate their potential. When the metatheory and its corresponding models and theories are poorly developed, examples become only tangentially related to them and difficult to evaluate vis-a-vis the theoretical or metatheoretical framework.

While the foregoing was meant to be illustrative of an ontological perspective to the development of complex educational programs, it does not necessarily depict the developmental sequence or the present status of Personalized Teacher Education. Perhaps, like most innovative programs, Personalized Teacher Education may have progressed from metatheory to example too quickly, without many of the intervening developments, i.e., theories and

models, which would assure a more well-defined and concrete example of the metatheory. For example, in mounting the current evaluation study, while admittedly the evaluation was conducted on a small-scale example of Personalized Teacher Education, sufficient modeling and theorizing, either conceptual or descriptive, had not been done to allow the evaluators to posit specific hypotheses for the evaluation study. The hypotheses stated in Chapter 2, the reader will recall, simply indicate that the evaluators and program developers had some general notions beforehand about the nature of interactions between student characteristics and training programs that they felt were communicated by the metatheory. Hypothesized behavioral outcomes as a function of specific instructional inputs could not be posited in the present study as much of the theoretical development and modeling related to Personalized Teacher Education remains to be done. In one sense, the current evaluation might have been premature inasmuch as its methodology assumed a fully developed concerns theory and model of personalization that, in actuality, are still undergoing development. It is interesting to note that, when examples of Personalized Teacher Education have been observed at other sites, it has often been difficult to discern the specific model and theories upon which a particular Personalized Teacher Education Program is based. While it has always been clear from such observations that the initial PTE metatheory provided the initial impetus for the program, it was not always clear that the theories and models employed by these programs fit the precise grammar and logic of the metatheory. This circumstance might be ascribed to the abstract or underdeveloped nature of the models and theories themselves or to on-site developers who have not operationalized these models and theories in sufficient detail. Such a circumstance is not new, of course, as it applies equally well to many of our most popular educational innovations. It is important that the development of personalized

theories and models continue so that the testing and evaluation of its examples are tied more closely to the metatheory. The Research and Development Center for Teacher Education has thus far developed a persuasive philosophy and metatheory of Personalized Teacher Education. It now may well be for others to further develop, demonstrate and evaluate the conceptual theories and models which embody this metatheory.

Some Comments About the Evaluation of Personalized Teacher Education

It has been apparent to those who have evaluated the personalized model of teacher training that there is a need for the concurrent development of the theory of personalization and methods for evaluating the theory. While the methods chosen to evaluate the Personalized Teacher Education Program generally have led to sound empirical studies, the usefulness and effectiveness of many other available methods have gone relatively unexplored. For a review of the broad range of methodologies that are available, evaluators should review Coan's (1968) analytical work on methods of inquiry in the social and behavioral sciences.

Coan, using an expansive analytical framework, factor analyzed ratings of outstanding psychological theorists concerning the various methodologies by which they investigated their theories. Coan's analysis identified six specific sets of methodologies with both quantitative and qualitative orientations. The quantitative methods which Coan identified emphasized such characteristics as observable behavior, operational definitions, statistical analysis, quantitative formulations and generalizability and generally were characterized by rigid control, reliance upon sensation and perception, and the investigation of immediate external determinants of behavior.

Coan's qualitative approaches pictured, on the other hand, a quite different set of methodological tools. These included such processes as

introspective reports, the clinical investigation of unconscious processes, naturalistic observation, evaluation of the uniqueness of the individual, and even armchair speculation. Coan characterized these latter, qualitative methodologies as a fluid orientation to the assessment of theory and the former, quantitative characteristics as a restrictive orientation. The latter, Coan concluded, provide a more general perspective on a theory's relationship to learning, motivation and affect than do the former.

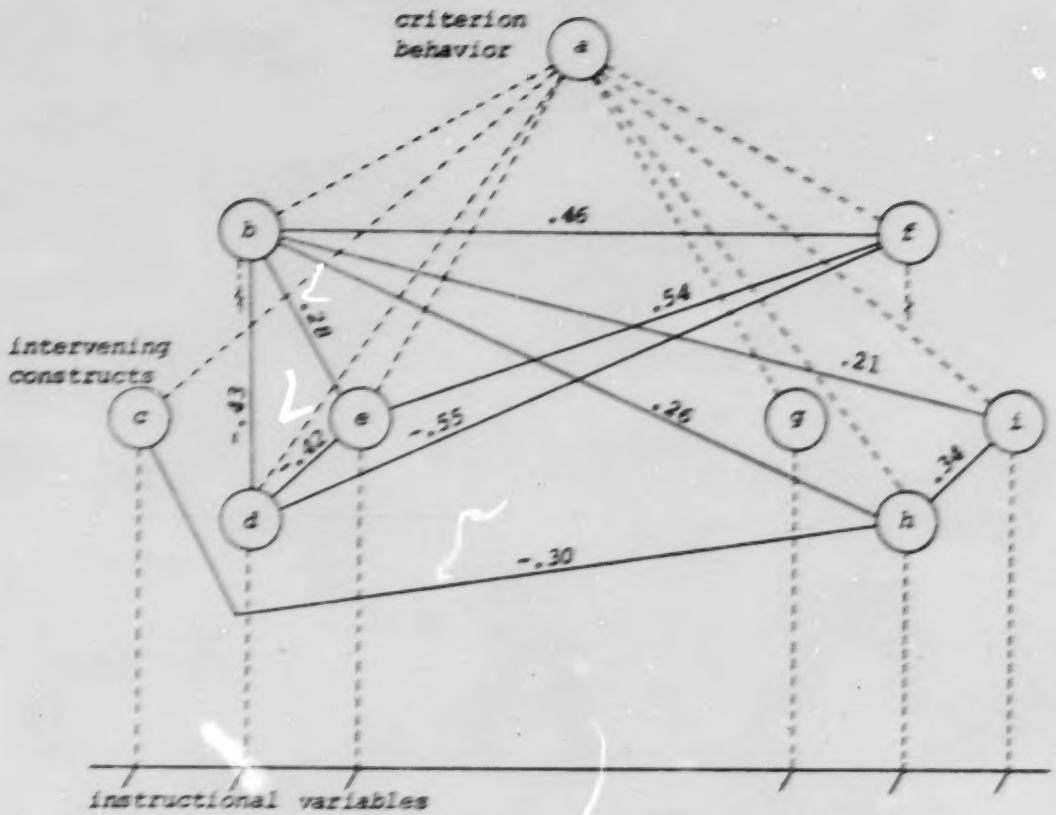
It would be fair to say that Coan's qualitative conceptualization is at odds with current methodological perspectives in evaluation, perspectives which rely heavily upon the traditional quantitative approach characterized by experimental control and, where possible, variable manipulation. Yet perhaps there is something to be learned from Coan's work in that when theory is not well defined and models ambiguous and contradictory, it is the fluid orientation to evaluation that is the most helpful to theory and model development. While clearly both qunatitative-restrictive approaches as well as qualitative-fluid ones are desirable, it was perhaps a limiting factor of the present evaluation that it concerned only those sets of behaviors that were most amenable to traditional statistical analyses which lie in the realm of the quantitative-restrictive approach. Armchair speculation, as Coan calls it, of unconscious processes leading to introspective reports based upon naturalistic--*in situ*--observation may well provide an equally and perhaps in some instances even more pervasive picture of how trainees perform and what trainees learn from a Personalized Teacher Education Program. In that the present evaluation study has been quantitatively oriented, it would be wise for future evaluations to employ more qualitative methodologies approaching more closely the criteria for naturalistic observation and hypothesis generation.

Evaluators of the personalized model of teacher training might utilize the criteria for evaluating its theories and models which have come to mind during the present evaluation. These criteria can be posed in the form of three questions: Is the metatheory of PTE useful? Are its theories and models truthful? Are its examples deployable to other settings and contexts?

Any approach to evaluating PTE should, first and foremost, strive to determine the usefulness of the undergirding metatheory of personalization. Tools and techniques of evaluation must be chosen to determine whether the general conceptualizations of intrapersonal, interpersonal and career-related behaviors and the interrelationships that have been posited to tie these concepts together is, in fact, a more effective approach to teacher training than many of the so-called traditional approaches. Second, any methodological tool must determine the truthfulness of personalized theories and models in depicting the real world. It is not uncommon for theories and models to schematize real-world events in so abstract a manner as for them not to be isomorphic with the real-world contingencies that exist within the context of a teacher training program. Methodologies must be suited to the ongoing training environment as opposed to the more isolated confines of the laboratory or simulated environment. Thirdly, any methodological approach to evaluating PTE must seek to determine the deployability of personalized theories and models in settings and contexts other than the one that has been used for the present evaluation study. These settings must include but should not be limited to competency-based programs and various other forms of traditional training. Any method chosen must bear the burden of determining the extent to which the concepts of Personalized Teacher Education are generalizable to other training programs and institutions.

These criteria suggest to evaluators a need to use both conceptual and descriptive methodologies for evaluating Personalized Teacher Education. Conceptual schemes, as will be noted below, can be employed as qualitative approaches to the evaluation of PTE simultaneous to the quantitative-empirical strategies of the nature reported in this volume. While quantitative strategies for evaluation have long been in use and are well documented in the evaluation literature, the following conceptual strategies, seemingly less used and infrequently reported, are also appropriate to the evaluation and development of Personalized Teacher Education. These conceptual strategies are (1) the construction and evaluation of nomological networks of observed, intervening and hypothetical behaviors, (2) the taxonomizing and sequential ordering of skills and competencies and (3) the use of path analytic tools to evaluate the causal sequences depicted in nomological networks and taxonomies.

Nomological networks. A potentially fruitful approach to positing and defining concepts within the Personalized Teacher Education Program is to hypothesize a sequence of intermediate or intervening behaviors that links terminal skills and competencies to instructional components of the training model. After reviewing psychological constructs related to terminal skills and competencies, the evaluator constructs what is called a nomological network of program behaviors. The evaluator posits causal connections between these psychological constructs and the terminal competencies and skills to be acquired. These "intervening" constructs, as shown in Figure 8-2, are related to observed data below and hypothetically to terminal skills and competencies above. The network of relationships becomes a working document with which the evaluator and developer revise and refine a program to bring about desired skills and competencies.



KEY TO CONSTRUCTS

(a) criterion, e.g., pupil gain
(b) self-confidence
(c) anxiety
(d) efficiency
(e) reality orientation
(f) sociability
(g) openness to feedback
(h) attitude toward children
(i) attitude toward authority

——— confirmed relationships, $p < .05$
- - - - - - - predicted relationships

Figure 8-2. Hypothesized network of constructs for the Personalized Teacher Education Program

A nomological network such as this could be used to substantiate the effects of intrapersonal and interpersonal behaviors upon career-related competencies and skills. Correlations and path coefficients could be computed to define relationships between observable data on the one hand and intervening constructs in a proposed theory or model of personalization on the other hand. For example, in the network shown above greater than chance correlations would be hypothesized between these intervening constructs and terminal skills and competencies, according to the proposed theory of personalization. Developers of Personalized Teacher Education should be especially interested in confirming relationships for which a small change in an intervening construct coincides with a large change in a desired competency or skill. These relationships then would be examined in experimental or quasi-experimental studies that would determine the extent to which intervening constructs are causal to career-related competencies and skills.

Nomological networks such as the one depicted above allow us to pass from intervening constructs to concrete skills and competencies and allow us to verify the existence of previously hypothetical behaviors. Two distinct types of activities can be used with the networks in establishing the presence of previously unconfirmed constructs. One activity occurs when the path that originally led to the formation of a construct or concept is retraced as would be illustrated if we moved from construct (a) to (b) to (d) in Figure 8-2. However, no new relationships are added to the definition of (a), the competency or skill, and so it is defined, in effect, by its antecedents. A second more informative type of reversal adds to the definition of the competency or skill by following a new path to other antecedents. The competency or skill becomes less tenuous as we substantiate relationships between it and other constructs. For example, relationships between competency (a)

and intervening constructs (b), (d) and (e) considerably expand the network of hypothesized causal agents and, in effect, the definition of the competency or skill itself. Relationships such as these can be documented and studied for the purpose of identifying the fluid or qualitative processes of which Coan speaks for the purpose of model and theory development. Borich and Drezek (1974), Yee and Gage (1968) and Duncan (1966) can provide the evaluator with the necessary background for measuring the presence and strength of causal paths.

Taxonomizing and sequential ordering of skills and competencies. A second potentially fruitful approach to defining relationships between performance criteria and intervening variables is that of taxonomizing and sequencing the intervening behaviors that the trainee is expected to acquire. Taxonomizing of the behaviors posited by the program's theories and models consists of a process of identifying a specific sequence in which the intervening constructs and competencies are acquired. For example, prospective teacher competencies and intrapersonal and interpersonal behaviors at various stages of the Personalized Teacher Education Program could be explicitly stated and the instructional components which generate these behaviors depicted as levels of expected knowledge in a taxonomy of the teacher training program. One might, for example, posit the sequence shown in Figure 8-3a, b. While highly schematized, these figures represent a series of specific instructional components for which there are one or more expressed behaviors in, let us say, either the intrapersonal, interpersonal or career-related domain. The question the evaluator must answer from such a taxonomy is whether the specified sequence of activities arranged by levels of intermediate knowledge is most appropriate for obtaining the expected competencies and skills at program completion. Let us indicate the attainment of any behavior with a plus sign and its

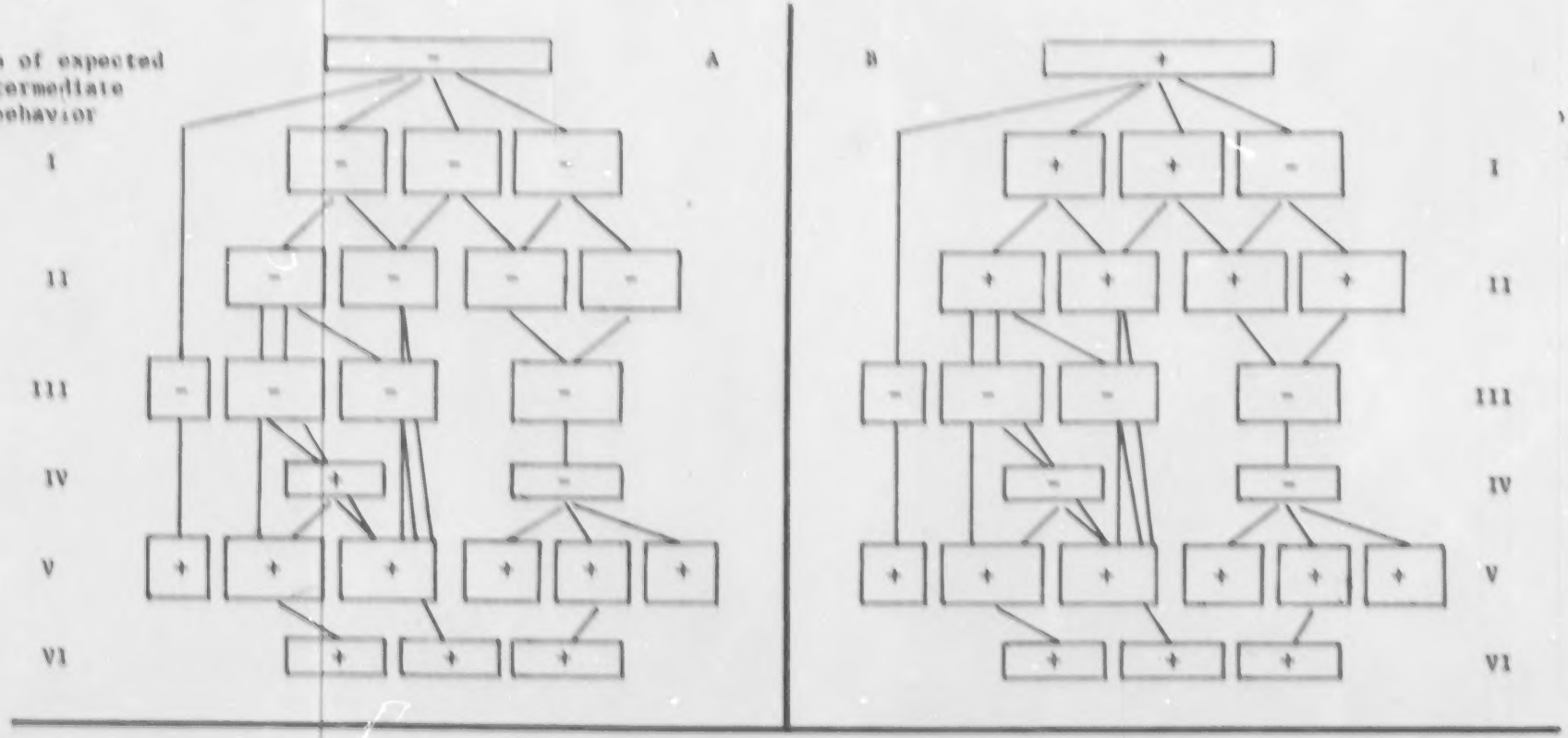


Figure 8-3a, b. Example of a pattern of attainment of learning sets in a hierarchy for two trainees, a low and a high achiever.

nonattainment with a minus sign. Moreover, let us assume Figures 8-3a and 8-3b represent the performance of two different trainees, one of whom reached criterion performance at program completion and one of whom did not.

For Teacher A it is clear that Level 3 behaviors represent a stumbling block to further attainment as no behavior more hierarchical in nature was acquired. While Teacher A has not attained the criterion behavior, the present sequence of instructional components may be reasonable, particularly if the program can provide additional instruction in smaller units between Levels 3 and 4.

For Trainee B, however, flaws in the basic conceptualization of what intermediate behavior is relevant to criterion performance and in the sequential arrangement of this behavior is indicated by the pattern of achievement. Although Teacher B achieved all expected behaviors at Levels 5 and 6, she failed to achieve any behaviors at Levels 4 and 5 and, most surprisingly, achieved behaviors at Levels 1 and 2, including the skills and competencies expected at program completion. Trainee B's pattern of achievement indicates a need to reconceptualize the instructional sequence, particularly the instructional components and behaviors expected at Levels 3 and 4.

Schematizations such as these can be used to represent the descriptive development of program theories and models while nomological networks of intervening and criterion variables can represent the conceptual development of these theories and models. The nomological network as well as the taxonomizing and sequencing of intervening behaviors and program competencies can contribute useful methodologies to the evaluation of the models and theories of personalization. Each of these approaches employs, in Coan's words, the fluid approach to evaluation as opposed to the more rigid constraints of statistical analysis which, unfortunately, has become synonymous with the quantitative orientation.

These statistical distinctions perhaps exaggerate to some extent the quantitative-qualitative continuum as both approaches could employ, as we have seen, sophisticated analytical techniques.

Some Concluding Technical Comments

It is perhaps appropriate in concluding this study to comment briefly upon the methodology that was employed. The reader will recall that two statistical approaches were taken to the evaluation of the personalized and traditional models of instruction. The first approach investigated main effects, i.e., mean differences between the personalized and traditional modes of teacher training across a variety of criterion variables, while the second approach investigated interactions between personality traits and attitudinal characteristics of the trainees and the training programs. While the analyses established that interactions do exist among these variables and programs, it is the purpose of the present section to reflect upon and to scrutinize the methodology that was used to find such interactions. Interwoven throughout these comments will be the authors' belief that an appropriate methodology for evaluating the Personalized Teacher Education Program is one that not only eliminates from consideration chance findings but also has the power to detect all of the significant findings that may be present.

The more general limitations of this study such as its sampling procedures (or lack of such), sample size and failure to hypothesize specific interactions are all too readily apparent and, therefore, will not be discussed. Instead, focus will be upon technical considerations that relate to the assignment of trainees to one program (e.g., personalized) or another program (e.g., conventional) on the basis of their entering personality and attitudinal characteristics. Let us review some obvious points first.

The interactions reported in this study offer little to the practitioner who wishes to assign trainees to competing programs on the basis of personality traits and attitudes. Far too many interactions and potentially relevant trait variables were reported in the current study to make the assignment of subjects to treatments feasible. The reader will recall that 22 trait variables revealed significant trait-treatment interactions across many criterion behaviors. Such a myriad of significant results precludes the development of any practical decision scheme that could be used to assign trainees to treatments on the basis of all of these findings. For example, a trainee who might have been assigned to personalized instruction on the basis of her pretest anxiety could also have been assigned to conventional instruction on the basis of her attitude-toward-children score. Assigning a trainee to a treatment on the basis of a single trait variable may, therefore, result in placing the trainee in the least effective treatment vis-a-vis some other trait variable. These contradictions are multiplied many times across individuals and across the 22 trait variables employed in this study.

The findings from this study should be viewed as heuristic in value in that they identify potentially important variables for which interactions between personality and attitude traits and training programs might exist. As the theoretical and conceptual framework of Personalized Teacher Education is developed, specific interaction hypotheses for these variables should be tested and, when found to relate to expected outcomes, can be used to assign students to treatments. Subsequent studies must focus upon specific predictor variables, i.e., intrapersonal and interpersonal behaviors, for which there is strong rationale vis-a-vis the metatheory.

It is important to note that regardless of the difficulty and inappropriateness of applying the results of this study to a practical setting, the

results that have been reported are, nevertheless, conservative in nature. The reader will recall the technique that was employed to rule out the reporting of chance events. The procedure required that greater than chance findings exist within an instrument across all criterion variables before any one analysis with that instrument was considered valid. When the number of significant findings for an instrument was not greater than chance, all findings for that instrument were ignored, even though some findings might have been potentially valuable ones. Even this procedure, however, led to far more significant results than could be interpreted.

It is important to note that even with these conservative procedures, there will remain an unreported level of error in the assignment of trainees to treatments on the basis of these results. The reader will recall that three statistical steps were completed in the calculation of trait-treatment interactions. The first step determined relationships between predictor and criterion variables within treatment groups. The second step involved testing the homogeneity of group regressions. And, the third step involved the calculation of regions of significance, i.e., ranges of predictor values for which one treatment was superior to the other.

The reader also will recall that most significant interactions revealed that the personalized model of teacher training was superior to the conventional model for trainees scoring at the low end of the predictor variable, i.e., in the left region of significance. Analogously, the conventional model was found superior to the personalized model for trainees who scored at the high end of the predictor variable, i.e., in the right region of significance. It would be deceiving to interpret these regions, however, as areas in which we are sure that the personalized model is superior to the conventional model or vice versa for all trainees, even though their calculation is a more conservative and commendable procedure than simply assigning trainees to treatments on the

basis of the predictor value at which the regression lines intersect. Even though we may be confident (if we have not violated any assumptions) that in 95 out of 100 studies conducted under the same conditions the boundary of the region of significance lies at approximately the same predictor value as was reported, we are not sure in any of these cases that a given personalized trainee in the left region of significance would, indeed, perform better than all conventional trainees in that region. To the contrary, some conventional trainees will perform better than personalized trainees even though the interaction and region of significance have indicated that the personalized model for a given predictor was superior to the conventional model. Figure 8-4 illustrates the problem.

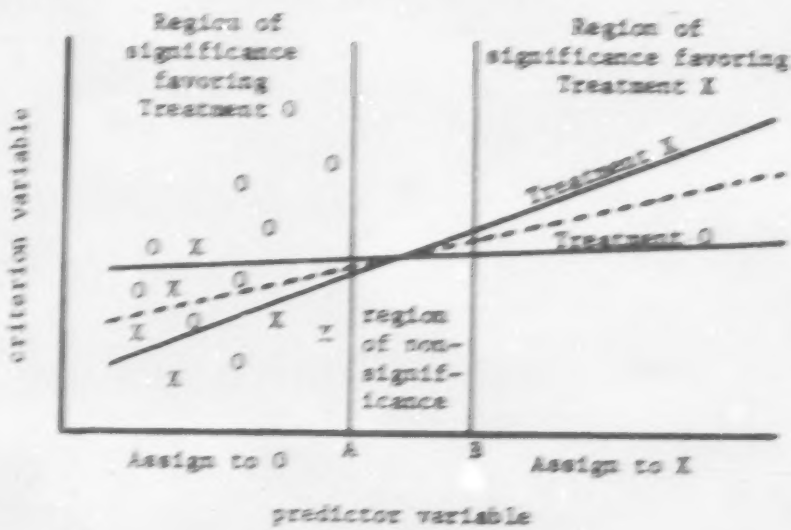


Figure 8-4. Residual error in the region of significance

Notice that, even though the personalized model of teacher training is shown to be superior to the conventional model for the area which lies to the left of the point at which the regressions intersect, some students in the conventional program (symbolized X) fall closer to the regression line for the personalized model of training than do some personalized students, and that some personalized students (symbolized O) fall closer to the regression line for the conventional training model than do some conventional students. We can expect this

overlapping to exist even when the investigator defines regions of significance with an acceptably high level of confidence.

Two approaches, one suggested by Borich (1973) and the other recently reported by Cronbach and Snow (1974), have been advanced as solutions to the problem. As neither of these appears in published form, their rationale and formulae are presented below.

Percent of error of assignment. Borich has proposed that researchers calculate what he has termed a percent of error in assignment index. This index is the percent of all subjects whose criterion score falls within a region of significance but who actually receive a criterion score inconsistent with their assigned group. This index is analogous to the number of subjects that a psychologist incorrectly categorizes or "misses" when a prediction formula is used to assign subjects to one of two groups.

The index is calculated by counting the number of subjects in the region of significance who, while assigned to the poorer treatment, actually performed above the midline between the regression lines for the two groups, i.e., a line equidistant from the two group regressions, and adding to this the number of subjects in the region of significance who, while assigned to the better treatment, actually performed below the midline between regressions. The percentage of both types of deviations within a region is calculated by finding the midline between the group regressions and then determining whether each observation falls above or below this line. These midlines are the dotted lines which appear in Figure 8-4. The midpoint for each subject at covariable X_1 is obtained by

$$Mpt_1 = \frac{\bar{Y}_1 + b_1(X_1 - \bar{X}_1) - \bar{Y}_2 - b_2(X_1 - \bar{X}_2)}{2} + \bar{Y}_2 + b_2(X_1 - \bar{X}_2)$$

or by the simplified equation:

$$Mpt_1 = \frac{\bar{Y}_1 + b_1(X_1 - \bar{X}_1) + \bar{Y}_2 + b_2(X_1 - \bar{X}_2)}{2}$$

where $\bar{Y}_1$, $\bar{X}_1$ and b_1 represent the criterion mean score, covariable mean score and regression coefficient, respectively, for the better treatment and $\bar{Y}_2$, $\bar{X}_2$ and b_2, these same values for the poorer treatment. The distance of each observation, y_1, from its respective midpoint is then given by

$$D = y_1 - Mpt_1.$$

D will be zero when the observation falls at the midpoint, positive when it falls above it, and negative when it falls below it. D's for observations assigned to the better treatment are expected to be positive and D's for observations assigned to the poorer treatment are expected to be negative. Exceptions are considered "misses" and are tallied and reported as a percent of the total number of observations within the region. For the data in Figure 8-4, two observations (0's) fell below the midline when they should have fallen above it and two observations (X's) fell above it when they should have fallen below it. Both types of deviations from the midline constitute 28 percent of the observations that lay within the region of significance. We, therefore, would report a 28-percent error in assignment if we chose to use the value "A" for assigning subjects who scored below this value to the better treatment.

It is important to note that, while this index takes into consideration the amount of "error" which can be expected about the group regressions, it does not provide information as to whether a subject has been assigned to a treatment incorrectly. This becomes obvious when we consider the case in which a subject who is assigned to the better treatment within a region of significance but whose score falls, let us say, at or below the regression for the

poorer treatment in this region is already performing the best that can be expected from either of the treatments. That is, by placing him in the opposing treatment we would depress his criterion score below even its present level. Therefore, while the residual error for a given subject may be large, it may, in fact, be less than that which would be encountered by placing the subject in the alternative treatment. Hence, while the percent of error in assignment index is an estimate of the overlapping cases that have occurred in a particular analysis, we cannot infer that the assignment of overlapping subjects to any other treatment would necessarily change their criterion performance.

The percent of error in assignment technique was not employed in this evaluation study but perhaps should have been. It is interesting to note, however, that in test applications of it, it was not uncommon to find examples in which regressions were heterogeneous and regions of significance definable that had a percent of error in assignment index as astonishingly high as 40 percent. That is, four out of every 10 subjects in the region of significance performed on the criterion measure more consistent with subjects in the opposing treatment than in their own treatment.

Simultaneous confidence interval. Cronbach and Snow have recently developed a second, albeit more complex, procedure with fewer of the interpretation problems inherent in Borich's technique. Cronbach and Snow's solution to the problem is to develop confidence intervals for the difference between regression lines at all values of the predictor variable. As Figure 8-5 illustrates, Cronbach and Snow's confidence region will be narrowest where group regressions intersect and largest where both Treatment A is better than Treatment B and where Treatment B is better than Treatment A. Cronbach and Snow's technique is essentially a confidence interval for the differences between means.

A direct statement about the limits of the interaction effect is attained by setting confidence limits on the population differences corresponding to the differences in outcome that describe the sample interaction. Such a confidence interval puts the differences between regression slopes for any given predictor variable in proper perspective in that we know that the observed differences are not the real differences as is shown by the hyperbola in Figure 8-5.

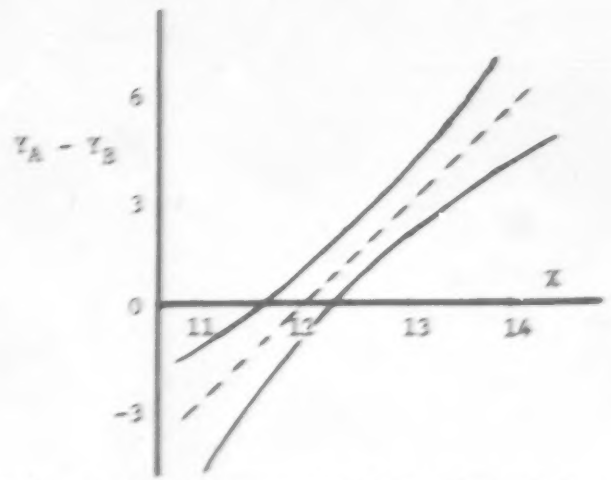


Figure 8-5. A simultaneous confidence interval around the difference $(Y'_A - Y'_B)$ in group regressions. Regressions intersect, i.e., $Y'_A - Y'_B = 0$, at $X = 12$.

For one predictor and two treatment groups of size N_a and N_b, the equation for this hyperbola is given by

$$\Delta = \left\{ 2F_{\alpha,d.f.} \left[\frac{1}{N_A} + \frac{1}{N_B} + \frac{(X-\bar{X}_{(A)})^2}{(N_A - 1)} \left(\frac{1}{s^2_{X(A)}} \right) + \frac{(X-\bar{X}_{(B)})^2}{N_B - 1} \left(\frac{1}{s^2_{X(B)}} \right) \right] s^2_e \right\}^{\frac{1}{2}}$$

where F is the value from the usual table where the intended confidence level is $1 - \alpha$ and d.f. is the number of degrees of freedom for SS Residual (here, equal to $N_A + N_B - 4$). The sample residual mean square, s^2_e, is the mean square of the deviations from the regression lines, pooled over treatments. The value of $\hat{Y}_A - \hat{Y}_B = \Delta\hat{Y}$ is obtained by subtracting one within-treatment regression

equation from the other. This function of X describes the interaction. $\Delta Y = \delta$, likewise a function of X, is the equation of the hyperbola that describes the confidence limits.

Cronbach and Snow refer to this calculation as a simultaneous confidence limit in that it is defining a confidence interval for all values of X. This approach is somewhat more conservative than the successive confidence interval noted by Potthoff (1964), as the latter will lead to a larger confidence interval and will fan out further toward both extremes of the distribution than will Cronbach and Snow's procedure. We now turn to one final methodological consideration, that of statistical power.

Statistical power. The term statistical power will be used to refer to the capacity of a statistical test to detect all the significant findings that are present. Statistical power is mathematically defined as $1 - \beta$, β being what is commonly referred to as a Type II or Beta error, i.e., the extent to which the investigator fails to reject a null hypothesis that is false or, simply, misses a significant effect when one is present. The reader will note that the above discussions of percent of error in assignment and, more directly, the estimation of confidence intervals for regression effects deal with problems typically associated with Type I errors or the probability of rejecting a null hypothesis that is true. Due to the number of significant findings reported in this study involving interactions, Type I errors are understandably of greater initial concern than are Type II errors. The issue of Type II errors in specific and of statistical power in general, however, is relevant to the present study and to the general problem of detecting significant results in evaluations of the personalized model of teacher training.

We have already mentioned a concern with the number of Type II errors that were likely to accrue from our attempts to report significant findings only when they were representative of an instrument for which 95 out of 100

analyses would be significant. We will now look at a somewhat related problem that suggests the use of a particular type of design that decreases the probability of missing significant interactions when they are present, thereby increasing statistical power.

The consideration of statistical power is crucial to any field of inquiry in which evaluators consistently fail to reject the null hypothesis. This has been somewhat the case in the field of trait-treatment interaction research wherein--contrary to the present study--relatively few such interactions have been reported. Bracht (1971) who conducted the field's most massive review of the literature to date reported that he could find only five significant interactions among 90 studies that hypothesized an aptitude by treatment interaction. Another review by Berliner and Cahen (1973), highlighting the conceptual and methodological problems of trait-treatment interaction research, offered a conclusion not unfamiliar to the readers of the Bracht article. This review suggested that both conceptual and methodological problems prevail in ATI research: conceptual problems related to designing studies that replicate and methodological problems related to finding interactions that may be present.

Cronbach and Snow (1973) have shown that for the case in which there is a moderately strong interaction the statistical power of the homogeneity of group regressions test is superior to blocking at the median, blocking at the 33rd and 67th percentiles or to blocking in any similar configuration that may be employed in a treatment by blocks design. Since classification schemes such as these discard power by treating dissimilar data as if they were the same, the degree of risk of an investigator's accepting a false null hypothesis is increased beyond that level which can be expected when the homogeneity of group regressions test is applied.

The evaluator who wishes to construct a more powerful design than can be attained with the homogeneity of group regressions test can, if his sample is sufficiently large, construct what is called an extreme groups design. The extreme groups design is constructed by dropping cases from the middle of the covariable distribution and by selecting the extreme cases from each tail of the distribution. In this manner, an appreciably more powerful design is constructed than with either a treatment by blocks design or by using the homogeneity of group regressions test that employs the full range of observed covariable values. Cronbach and Snow (1973) have concluded that by cutting at the quartile points only about two-thirds of the cases needed for the homogeneity of group regressions test are required to maintain the same level of power with an extreme groups design. With more extreme cuts, even greater efficiency can be obtained.

The standard analysis for the extremes groups design is a treatment by levels analysis of variance. Covariable scores are usually classified into high and low categories, each containing equal numbers of cases. The analysis yields mean squares for the treatments, levels, treatment by levels, and residual error. The F-ratio for treatments by levels is the test for interaction. Extreme groups designs are discussed further by Borich and Godbout (1974) and are recommended for the study of trait by training program interactions in the evaluation of Personalized Teacher Education.

This chapter has reviewed the conceptual development of Personalized Teacher Education and has identified several formative strategies for evaluating its theories and models. Three central points were made in the context of this discussion.

1. It was suggested that, while the Personalized Teacher Education literature communicates a persuasive metatheory, its theories and models are not well developed and not clearly communicated. Underdeveloped theories and models, it was contended, lead to the construction of weak program examples that are tied only tangentially to the metatheory.

2. As a result of this observation the authors suggested that such tools as nomological networks and taxonomies of sequentially ordered behavior which are representative of a fluid orientation to the evaluation of personalized training are more appropriate at present than the more restrictive strategies embodied by traditional statistical designs.

3. Lastly, it was suggested that future evaluations of personalized training employ methodologies that test the usefulness of the metatheory, the truthfulness of its theories and models, and the deployability of its examples. In conjunction with these criteria, it was suggested that methodologies be chosen on the basis of their capacity to reject significant findings that are due to chance as well as on the basis of their capacity to detect all significant findings that are present. Calculations for the percent of error in assignment index and the simultaneous confidence interval were provided as techniques for rejecting significant findings due to chance and the extreme groups design discussed as a technique for increasing the statistical power of an evaluation design, i.e., the likelihood of detecting all the significant findings that are present.

Appendix A

GUIDELINES FOR STUDENT TEACHERS AND PUBLIC SCHOOL SUPERVISING TEACHERS CONCERNING VIDEOTAPING
(Pretest)

Introduction. Approximately two weeks from today student teachers at your elementary school will teach a 20-minute mini-lesson to 10 pupils. The lesson will be videotaped. This videotape is not to show each student teacher at her best, but rather to capture a representative sample of her teaching.

Student teachers from several exemplary teacher training programs will be making such videotapes at the beginning and end of this semester's student teaching experience. The purpose is simply to see if students from different programs teach differently, and how their teaching changes over the course of student teaching.

This is not a test. Individual student teachers are not being evaluated. Their videotapes will not go to the College of Education.

The Videotaped Lesson. Each student teacher will write one to three objectives for her learners, relative to the following general goal:

> The student teacher will introduce a set of terms (e.g. concepts, symbols, technical terms...) which her pupils probably have not encountered previously. The terms may be drawn from the subject areas of Language Arts, Reading, Social Studies or Science. The set of terms introduced should become part of the working knowledge of each learner.

Note that a large amount of teaching is just this -- conveying a meaning for symbols such that learners can operate in the world. Preparation should include selecting for your pupils content which will supplement their current classroom work. Consulting with supervising teachers is a good idea.

There is no correct way to teach the lesson. You need not evaluate your objectives (there is not enough time), and your students will not be tested on them. Simply teach this lesson as you would any other.

Preparation. How long is this lesson? Twenty minutes. That is less than 10% of a teaching day. When you're teaching full-time how long can you spend actually preparing for the next day? Probably a couple of hours at most. It would be unreasonable, then, to spend a great deal of time preparing this lesson. We suggest an hour or less; treat this lesson as an important lesson for the day you actually teach it.

Materials. You may use whatever supplies you wish to bring to the videotaping session.

Physical Organization. You may arrange the pupils and room as you wish, given the limitations of camera placement. The crew will assist you.

Scheduling. You can sign-up for a time on the schedule posted by your college supervising teacher. We are videotaping nearly 100 student teachers all across the city, so be prompt. Also please exercise care in negotiating a time convenient for your public school supervising teacher to release 10 pupils. We are on a 35-minute schedule, which should allow you time to get settled and begin gracefully.

Selecting Pupils. Public school supervising teachers (not student teachers) are asked to select 10 pupils representative of the pupils a given student teacher is currently teaching. The fairest way to do this is to select a random group from a larger group of pupils the student teacher comes in contact with. We recognize that in a given school with its scheduling problems it may not be easy to schedule pupils; therefore, please follow one of these options. Indicate the one you followed on the last page where you list pupils.

Option 1. Use a homeroom list or other administrative grouping.
(If this is not possible)

Option 2. Use a class list e.g. Language Arts or PE.
(If this is not possible)

Option 3. Have the student teacher list herself the pupils she currently has contact with and use this list.
(If this is not possible)

Option 4. Use a group of pupils the student teacher currently teaches as a natural group e.g. a reading group. However, do not choose a high ability or low ability grouping; choose a middle ability group, if your school ability groups.

Referring to the list of possible pupils (option 1, 2, or 3),

(1) Number all pupils consecutively. If you have a roll divided into sexes, be careful to use a true alphabetic order, numbering without regard to sex.

(2) Send to the videotaping whoever is numbered 3, 7, 8, 9, 10, 13, 14, 17, 18, and 21. (These were randomly selected numbers).

(3) If students with one of the above numbers are not appropriate (e.g., hard-of-hearing or parents strongly object) send number 1, 5, 11, 12, and 15 in that order of replacement.

Mechanics of the Lesson. One member of the videotaping crew will take you smoothly and naturally through the following steps. You will:

1. Bring your pupils to the room a few minutes early.

2. Arrange the pupils and room.

3. Announce to the cameraman that you are ready to begin (a microphone will be placed around your neck at this time).

4. Teach your lesson for 20 minutes (clock provided), announcing to the cameraman when you are through.

   (Because of the tight schedule, at exactly 20 minutes the videotape camera will stop, the crew will indicate this, and you will have a couple of minutes to end the lesson naturally. Then, if you are still not finished, the crew will have to interrupt you.)

5. Give your pupils' attention to the crew member in charge.

6. Complete a brief (2 minutes) questionnaire on how natural and representative you felt your teaching was, while the pupils do the same.

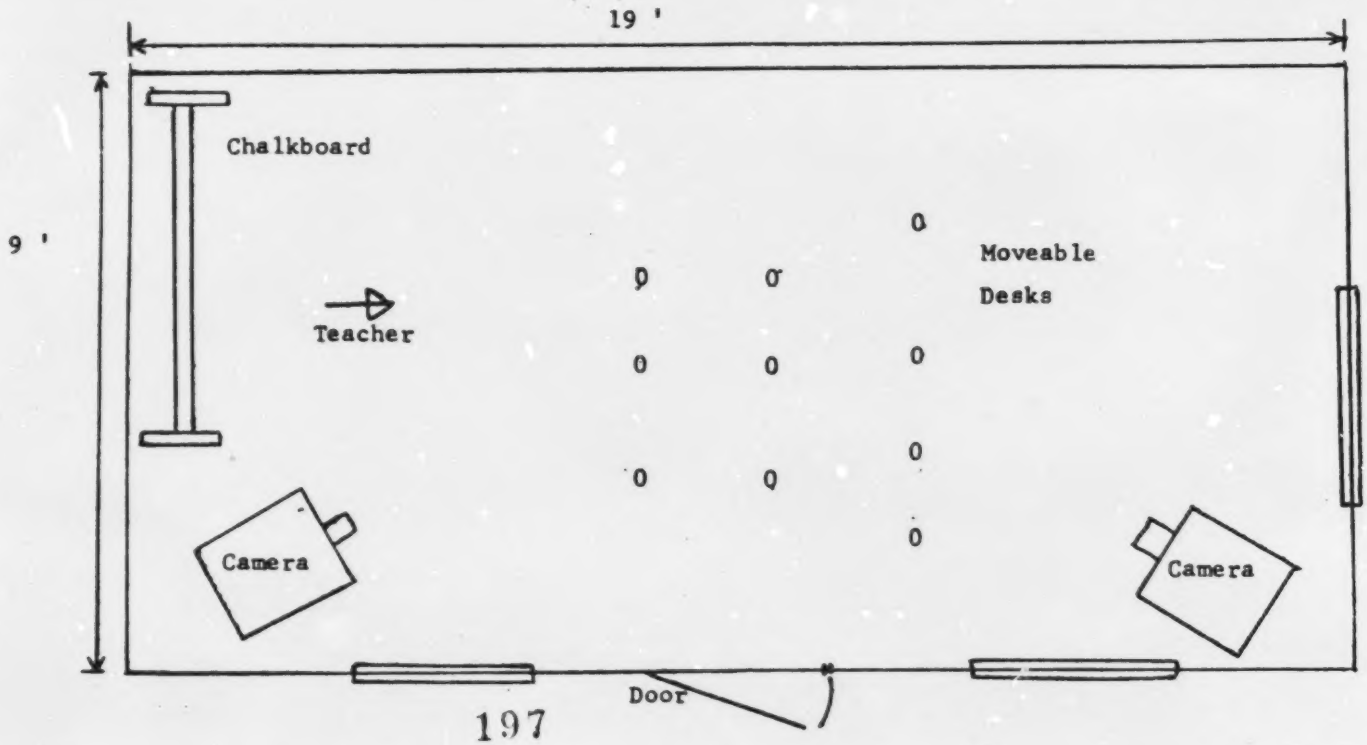
7. Return your pupils.

After consulting with your public school supervising teacher please complete the following page to be handed to the videotape crew summarizing the mechanics of this lesson.

If there is anything you have a question about with regard to the videotaping feel free to call:

471-1209 (office)
472-5325 (home)

Portable buildings. While a regular classroom is free for videotaping at School 3, at other schools portable buildings will be used. Their floor plan follows:



Appendix A (cont.)

Do not write in this space

I.D.#_____

Student Teacher_____

School _____

The following pupils have been assigned to me by my public school supervising teacher as representative of the pupils I am currently teaching.

_____ _____

_____ _____

_____ _____

_____ _____

"I understand the purpose of this videotaping, have selected the pupils listed above, and have planned on the time scheduled below."

Signature of Public School Supervising Teacher_____

Scheduled starting time_____ : _____ / _____ / _____
(hour) (day) (month)

Student teacher's objectives for this lesson:

1._____

_____

_____

2._____

_____

_____

3._____

_____

_____

Appendix A (cont.)

GUIDELINES FOR STUDENT TEACHERS AND
PUBLIC SCHOOL SUPERVISING TEACHERS
FOR SECOND VIDEOTAPING SESSION
(Posttest)

Second Videotaping Session. In approximately two weeks the Videotaping crew will be with you again to tape your second 20-minute mini-lesson session. We'd like to share with you some knowledge gained from the first videotaping session to help you prepare for this second session and also to reemphasize that THIS IS NOT A TEST. Although it is understandable that each student teacher likes to do her very best every time she teaches, the purpose of this videotape is not to show each student teacher at her best but rather to capture, as nearly as these circumstances will allow, a representative sample of her teaching at that particular point in her professional program. Individual student teachers are not being evaluated nor will their videotapes go to the College of Education. The purpose of these videotapes is to collect a sufficiently large body of data to see if students from different programs teach differently and how their teaching changes over the course of student teaching. The videotapes will be coded to gain this information. Those student teachers for whom videotaping forms a part of their professional training program know that in some cases their videotapes will be reviewed by their counselors and themselves. All of you know that your tape is available for your own review if you care to come in and see it. But in all cases, these tapes are CONFIDENTIAL DATA and are not shown to anyone without your permission.

The Videotaped Lesson. Since you have already experienced one videotaping session you are now familiar with the process. For the stability of this study we must ask that certain things be kept constant, and we will appreciate your following these guidelines when preparing for your second videotaping session.

Length of Lesson: This videotaped lesson will again be 20 minutes long.

Lesson Content: It is crucial to this study that lesson content be kept constant. This means that we are again asking you to prepare your lesson from the subject areas of Language Arts, Reading, Social Studies, or Science, and to introduce or present concepts, terms, or skills which you believe will be new to your pupils. We would appreciate it if you would avoid a strictly "review" lesson or one that deals only with the application of a previously taught skill. Please do not interpret this to mean that if your first lesson was in the subject area of Reading you must teach this second lesson in the same subject area. We are simply asking that you stay within the subject areas and purposes listed above.

Preparation: We will again ask you to state the objectives for your lesson on the attached sheet which is to be given to the Videotaping crew just prior to your taping session. We suggest that three objectives are a maximum for a 20 minute lesson and also that you need not concern yourself with an evaluation objective within this lesson period. Your

School Supervising Teacher is undoubtedly your best resource in planning subject content for this session, along with your own previous experience. Do not prepare excessively--remember, the point of the filming is to record typical, not ideal, performance.

Materials and Physical Organization: With the exception that we will be using an area rug and will not provide tables or seats for your pupils, your physical circumstances for this second session will be exactly the same as the first. If you plan to have your pupils do any writing, you will need to bring something for use as a writing surface.

Selecting Pupils. Public school supervising teachers are asked to select pupils representative of the pupils a given student teacher is currently teaching. With our first videotaping experience behind us we are now asking that you bring eight (8) pupils rather than ten to the videotaping session. The fairest way to select pupils is to pick a random group from a larger group of pupils with whom the student teacher comes in contact. This is probably a homeroom group, record-keeping group, or Language Arts group. You decide on the appropriate larger group to consider. Then, from such a group, please take the roll of students and number all pupils consecutively. (If you have a roll divided into sexes, be careful to use a true alphabetic order, numbering without regard to sex.) Then send to the videotaping pupils numbered 3, 7, 8, 9, 10, 13, 14, and 17. (These were randomly selected numbers.) If students with one of the above numbers are not appropriate (e.g., hard-of-hearing, or parents would strongly object, etc.) numbers 1, 5, 11, 12, and 15 can be substituted (in that order of replacement, please). Using the numbers above should assure that each student teacher is treated equally with respect to the pupils she is to teach.

Pupils in More Than One Session. For this second session not only the student teacher but most of her pupils will have had previous experience with videotaping. This will be pretty much the general situation and should not be a point of concern in selecting pupils. We can say, however, that pupils who are included in several sessions on the same day do have a tendency to become bored old-timers and seem pressed to find some way to keep the experience novel. If possible, it might be a good idea to check pupil listings with each other.

Mechanics of the Videotaped Lesson. One member of the videotaping crew will take you smoothly and naturally through the following steps. You are asked to:

1. Bring your pupils to the room a few minutes early.
2. Give a crew member your Student Listing and Objectives Form.
3. Arrange (with assistance from the crew) the pupils and room.
4. Announce to the cameraman that you are ready to begin. (A mike will be placed around your neck at this time.)
5. Teach your lesson for 20 minutes, announcing to the cameraman when you are through. You will be provided with time signals.
6. Direct your pupils' attention to the crew member in charge.

7. Complete a brief (2 minutes) questionnaire on how natural and representative you felt your teaching was.

8. Return your pupils to class.

If there is anything you have a question about with regard to the videotaping, please feel free to call:

471-1209 (Office)
472-5325 (Home)

Scheduling Your Videotaping Lesson Time. The time scheduling process for this second videotaping session will be conducted in the same manner as the first session. When you have selected your videotaping time, make a note of that time for your own information. In our first videotaping sessions we did find that it was possible for student teachers to forget or confuse their appointment times and dates.

Thank you.

Appendix A (Cont.)

Do not write in this space

I.D.#_____

Student Teacher_____

School_____

The following pupils have been assigned to me by my public school supervising teacher as representative of the pupils I am currently teaching.

_____ _____

_____ _____

_____ _____

_____ _____

"I understand the purpose of this videotaping, have selected the pupils listed above, and have planned on the time scheduled below."

Signature of Public School Supervising Teacher_____

Scheduled starting time_____ : _____ / _____ / _____
(hour) (day) (month)

Student teacher's objectives for this lesson:

1._____

_____

_____

2._____

_____

_____

3._____

_____ _____

_____

Appendix B

Significant Trait-Treatment Interaction Results

This appendix presents the 107 significant trait-treatment interaction results which were accepted as valid. The "Interaction" column presents the trait variable involved in the specific trait-treatment interaction. The probability ("Prob."), "F," "df," %variance ("%Var."), and beta-weight ("Beta") columns are relevant to the homogeniety of group regressions test. "Beta" involves the criterion-on-trait regression seperately for the "PTE" and non-PTE("Non") programs. "Point of Intersect" refers to the point of intersection between the PTE criterion-on-trait regression line and the non-PTE criterion-on-trait regression line. The "Lower Region of significance" and "Upper Region of significance" columns present the results of the Johnson-Neyman analyses. The "bounds" columns refer to the boundaries of the region in question and "%subjects" refers to the percentage of the total sample of subjects falling within that region. Finally, the "Correl." (correlation) column is relevant to the relationship between the trait variable and the criterion variable seperately for the "PTE" program and for the "Non" (non-PTE) program.

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
For Criterion Variable: Student Evaluation of Teacher Training Program: Behavior Modeling											
ASD/ Efficiency	.02	5.40	1/59	8.29	30.45	8 to 20.29	0	35.79 to 40	16	PTE: -1.07 NON: 1.06	PTE: - .29 NON: .33
ASD/ Introversion	.00	11.54	1/59	16.21	22.49	8 to 17.88	25	28.28 to 40	14	PTE: .87 NON: - .48	PTE: .42 NON: - .39
ASD/ Attractiveness	.01	6.54	1/59	8.80	24.93	8 to 17.36	2	29.72 to 40	23	PTE: -1.53 NON: .11	PTE: - .54 NON: .07
SRI/ Self	.07	3.24	1/59	5.11	19.50					PTE: - .78 NON: .46	PTE: - .28 NON: .20
SRI/ Others	.03	4.80	1/59	7.45	19.52	6 to 8.80	0	24.25 to 30	0	PTE: -1.10 NON: .69	PTE: - .29 NON: .27
SRI/ Children	.08	3.15	1/59	4.99	21.87					PTE: - .20 NON: 2.29	PTE: - .06 NON: .42
SRI/ Authority	.01	6.25	1/59	9.53	17.88	6 to 11.57	3	21.88 to 30	20	PTE: - .84 NON: .79	PTE: - .27 NON: .38

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	Lower Region of significance %subjects	Upper Region of significance bounds	Upper Region of significance %subjects	Beta	Correl.
SRI/Hope	.02	2.82	1/59	4.14	19.49					PTE: -1.50 NON: - .02	PTE: - .43 NON: - .04
SRI/ Work	.00	8.16	1/59	11.93	15.95	6 to 11.15	9	19.48 to 30	30	PTE: - .82 NON: .66	PTE: - .37 NON: .36
SRI/ Reality	.03	4.92	1/59	7.33	15.45	6 to 1.95	0	20.19 to 30	9	PTE: - .90 NON: .34	PTE: - .40 NON: .18
OWSC/ Evasion	.00	8.62	1/62	11.99	1.87			4.22 to 62	9	PTE: - .30 NON: .29	PTE: - .34 NON: .41
OWSC/ Depression	.07	3.33		4.96	1.89					PTE: .31 NON: - .12	PTE: .34 NON: - .15
For Criterion Variable: Student Evaluation of Teacher Training Program: Person-Centered											
ASD/ Attitude	.02	5.97	1/57	9.12	33.06	8 to 21.95	1	36.44 to 40	38	PTE: -1.29 NON: 1.47	PTE: - .34 NON: .31
ASD/ Attractiveness	.01	6.45	1/57	4.58	22.38			7.30 to 62	23	PTE: - .27 NON: .92	PTE: - .11 NON: - .35

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
OWSC/ Depression	.01	6.45	1/60	9.35	2.73	0 to 1.43	46	7.30 to 62	2	PTE: .26 NON: -.29	PTE: .29 NON: -.35
For criterion variable: Student Evaluation of Teacher Training Program: Personal, Intellectual and Social Development											
ASD/ Behavior	.01	6.48	1/58	9.91	13.10	8 to 8.09	1	17.59 to 40	14	PTE: -.15 NON: 1.02	PTE: -.10 NON: .46
SRI/ Self	.01	6.46	1/58	9.89	20.82	6 to 16.08	9	25.93 to 30	1	PTE: .32 NON: -1.43	PTE: .15 NON: -.44
SRI/Work	.02	5.43	1/58	8.47	16.56	6 to 9.32	4	23.26 to 30	6	PTE: .24 NON: -.98	PTE: .24 NON: -.41
SRI/ Reality	.03	5.04	1/58	7.28	16.72	6 to 10.32	5	25.11 to 30	1	PTE: -.20 NON: -1.49	PTE: -.12 NON: -.55
For Criterion Variable: Set 2: Rapport											
ASD/ Attitude	.04	4.41	1/62	6.59	34.61	8 to 19.09	1			PTE: .33 NON: -2.03	PTE: .09 NON: -.47

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
ASD/ Behavior	.07	3.28	1/62	4.96	13.59					PTE: - .06 NON: .76	PTE: .04 NON: .46
ASD / Efficiency	.09	2.83	1/62	4.34	30.80					PTE: .37 NON: -1.15	PTE: .11 NON: - .36
ASD/ Introversion	.07	3.43	1/62	5.20	22.26					PTE: - .58 NON: .19	PTE: - .26 NON: .16
SRI/ Self	.01	6.41	1/62	9.29	20.67	6 to 15.64	8	25.42 to 30	1	PTE: .36 NON: -1.37	PTE: .14 NON: - .55
SRI/ Others	.05	3.82	1/62	5.77	19.79					PTE: .47 NON: -1.13	PTE: .14 NON: - .41
SRI/ Children	.03	4.92	1/62	7.28	22.29	6 to 17.88	5	26.03 to 30	0	PTE: .38 NON: -2.69	PTE: .09 NON: - .50
SRI/ Hope	.03	4.59	1/62	6.71	20.31	6 to 9.49	1	26.08 to 30	0	PTE: 1.24 NON: - .51	PTE: .34 NON: - .20
OWSC/ Depression	.02	6.10	1/64	8.70	1.96			4.55 to 62	3	PTE: - .23 NON: .30	PTE: - .22 NON: .45

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	Lower Region of significance %subjects	Upper Region of significance bounds	Upper Region of significance %subjects	Beta	Correl.
For Criterion Variable: Set 2: Unreasonable Negativity											
ASD/ Attractiveness	.10	2.78	1/62	4.22	24.52					PTE: -.71 NON: .48	PTE: -.24 NON: .20
SRI/ Others	.09	2.95	1/62	4.50	19.17					PTE: -.43 NON: .98	PTE: -.13 NON: .34
SRI/ Authority	.09	2.84	1/62	4.25	17.89					PTE: -.96 NON: .13	PTE: -.31 NON: .08
For Criterion Variable: Set 2: Posterance of Self-Esteem											
ASD/ Attitude	.09	2.87	1/62	4.36	36.49					PTE: .29 NON: -1.63	PTE: .10 NON: -.30
SRI/ Others	.03	4.68	1/62	6.95	21.13	6 to 17.37	10			PTE: .95 NON: -.83	PTE: .28 NON: -.26
OWSC/ Response Length	.08	3.08	1/64	4.30	6.46					PTE: -2.94 NON: 0	PTE: -.43 NON: .02

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
OWSC/ Populars	.09	2.95	1/64	4.23	20.31					PTE: .81 NON: - .07	PTE: .37 NON: - .05
OWSC/ Evasion	.06	3.56	1/64	5.24	2.55					PTE: .23 NON: - .14	PTE: .26 NON: .19
For Criterion Variable: Readiness Assessment (College Supervisor): Self Concern											
ASD/ Attitude	.01	7.31	1/62	10.48	35.29	8 to 31.44	13	39.81 to 40	7	PTE: -1.17 NON: 1.84	PTE: - .30 NON: .36
ASD/ Behavior	.00	11.78	1/62	15.32	12.95	8 to 9.94	10	15.49 to 40	30	PTE: .12 NON: -1.32	PTE: .08 NON: - .65
ASD/ Efficiency	.00	10.21	1/62	13.98	31.64	8 to 28.25	26	35.23 to 40	16	PTE: - .67 NON: 2.03	PTE: - .21 NON: .53
SRI/ Others	.00	8.99	1/62	12.58	20.31	6 to 17.76	10	23.19 to 30	13	PTE: -1.37 NON: 1.04	PTE: - .39 NON: .32
SRI/ Children	.01	7.07	1/62	9.02	22.77	6 to 20.98	18	25.64 to 30	0	PTE: .88 NON: 4.24	PTE: .17 NON: .65

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	Lower Region subjects	Upper Region of significance bounds	Upper Region subjects	Beta	Correl.
SRI/Authority	.04	4.13	1/62	6.14	18.62					PTE: - .26 NON: .99	PTE: - .11 NON: .39
SRI/Work	.00	16.55	1/62	20.82	16.95	6 to 14.63	30	19.42 to 30	30	PTE: - .56 NON: 1.35	PTE: - .29 NON: .62
OWSC/Response Length	.09	2.87	1/64	4.16	6.52					PTE: .30 NON: -2.49	PTE: .08 NON: - .33
For Criterion Variable: Readiness Assessment (College Supervisor): Concern for Children											
ASD/Attitude	.07	3.41	1/62	5.13	36.49					PTE: - .90 NON: 1.20	PTE: - .23 NON: .24
ASD/Behavior	.02	6.18	1/62	8.94	12.18			15.82 to 40	30	PTE: .22 NON: .89	PTE: .14 NON: - .43
ASD/Efficiency	.08	3.14	1/62	4.62	33.16					PTE: - .34 NON: 1.21	PTE: - .11 NON: .31
SRI/Work	.03	4.87	1/62	7.19	18.13	6 to 13.03	21			PTE: - .30 NON: .82	PTE: - .16 NON: .38

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	Lower Region of significance %subjects	Upper Region of significance bounds	Upper Region of significance %subjects	Beta	Correl.
OWSC/ Populars	.05	3.79	1/64	5.36	19.28	0 to 6.81	1			PTE: - .05 NON: .90	PTE: - .03 NON: .44
	For Criterion Variable: Readiness Assessment (College Supervisor): Concern For Impact										
ASD/ Attitude	.04	4.16	1/62	6.13	35.96	8 to 29.02	5			PTE: -1.28 NON: .98	PTE: - .32 NON: .22
ASD/ Behavior	.01	7.70	1/62	10.75	12.53			15.78	30	PTE: .12 NON: -1.08	PTE: .08 NON: - .56
ASD/ Efficiency	.04	4.44	1/62	6.64	32.36	8 to 24.91	12			PTE: - .58 NON: 1.28	PTE: - .18 NON: .35
SRI/ Others	.03	4.88	1/62	7.26	20.74	6 to 16.45	7			PTE: - .92 NON: .91	PTE: - .25 NON: .30
SRI/ Children	.10	2.83	1/62	4.06	23.18					PTE: .68 NON: 2.89	PTE: .12 NON: .47
SRI/ Work	.00	9.41	1/62	13.10	17.40	6 to 14.18	30	21.62 to 30	21	PTE: - .50 NON: .91	PTE: - .25 NON: .30

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	% subjects	Upper Region of significance bounds	% subjects	Beta	Correl.
SRI/ Reality	.09	2.97	1/62	4.40	17.01					PTE: - .75 NON: .20	PTE: - .35 NON: .10
SRI/ Hope	.07	3.39	1/62	5.11	21.34					PTE: - .96 NON: .58	PTE: - .28 NON: .18
OWSC/ Response Length	.07	3.46	1/64	5.05	6.46					PTE: 1.13 NON: -2.12	PTE: .18 NON: - .28
For Criterion Variable: Readiness Assessment (College Supervisor): Perceptive about Self											
ASD/ Efficiency	.09	2.86	1/63	4.26	31.95					PTE: -1.02 NON: .44	PTE: - .27 NON: .16
SRI/ Others	.04	4.33	1/63	6.40	20.49	6 to 13.21	5			PTE: - .96 NON: .76	PTE: - .25 NON: .28
SRI/ Authority	.06	3.61	1/63	5.38	19.15					PTE: - .81 NON: .43	PTE: - .25 NON: .21
SRI/ Work	.02	8.57	1/63	7.82	17.11	6 to 11.16	9	24.76 to 30	0	PTE: - .89 NON: .28	PTE: .40 NON: .17

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
SRI/ Hope	.02	6.17	1/63	8.88	20.96	6 to 17.20	14	25.49 to 30	0	PTE: -1.01 NON: 1.06	PTE: - .27 NON: .36
OWSC/ Response Length	.03	4.57	1/65	6.36	6.46	0 to 3.44	0	7.01 to 62	15	PTE: 2.52 NON: - .98	PTE: .33 NON: - .18

For Criterion Variable: Readiness Assessment (College Supervisor): Perceptive about Children

ASD/ Attitude	.06	3.48	1/61	5.17	38.27	8 to 34.08	37			PTE: - .74 NON: 1.40	PTE: - .22 NON: .24
ASD/ Efficiency	.07	3.43	1/61	5.11	35.16	8 to 30.30	38			PTE: - .58 NON: 1.08	PTE: - .19 NON: .27
SRI/ Others	.02	5.19	1/61	7.52	22.34	6 to 19.62	35			PTE: - .98 NON: .87	PTE: - .30 NON: .26
SRI/ Children	.07	3.33	1/61	4.75	24.31	3.14 to 22.15	37			PTE: .40 NON: 2.84	PTE: .07 NON: .42
SRI/ Work	.01	6.38	1/61	9.09	19.57	6 to 16.15	53			PTE: - .63 NON: .64	PTE: - .33 NON: .29

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
OWSC/ Response Length	.03	4.98	1/63	6.71	6.21	0 to 2.31	0	6.62 to 10	49	PTE: .52 NON: -3.11	PTE: .11 NON: - .41

For Criterion Variable: Fair: Teacher is Tangential

ASD/ Attitude	.07	3.28	1/51	5.73	36.63					PTE: .23 NON: -2.25	PTE: .06 NON: - .54
ASD/ Attractiveness	.01	7.87	1/51	13.02	27.36	8 to 23.54	25	34.53 to 40	3	PTE: .42 NON: -1.55	PTE: .16 NON: - .65
SRI/ Others	.01	7.64	1/51	12.85	21.00	6 to 18.36	23	25.92 to 30	0	PTE: 1.01 NON: - .86	PTE: .29 NON: - .49
SRI/ Work	.02	5.92	1/51	10.25	18.03	6 to 13.93	21	28.61 to 30	0	PTE: .46 NON: - .86	PTE: .22 NON: - .49
OWSC/ Populars	.09	2.91	1/54	4.97	21.42	5.26 to 13.92	16			PTE: .64 NON: - .29	PTE: .27 NON: - .19
OWSC/ Depression	.09	2.99	1/54	5.10	0.66			3.38 to 6.44	8	PTE: - .06 NON: .30	PTE: - .06 NON: .43

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
For Criterion Variable: Fair: Student Explores											
SRI/ Children	.07	3.37	1/51	6.07	21.27			25.38 to 29.13	0	PTE: -1.27 NON: 1.36	PTE: - .44 NON: .18
OWSC/ Populars	.00	12.88	1/54	16.37	19.46	0 to 16.67	39	23.44 to 62	12	PTE: .00 NON: -1.57	PTE: .03 NON: - .56
OWSC/ Evasion	.00	10.83	1/54	9.41	1.12			2.74 to 62	27	PTE: - .02 NON: .77	PTE: - .08 NON: .75
For Criterion Variable: Fair: Student Questions											
ASD/ Attractiveness	.02	5.93	1/51	10.08	27.26	8 to 22.39	21	38.44 to 40	0	PTE: - .22 NON: 1.51	PTE: - .10 NON: .51
SRI/ Work	.03	4.64	1/51	8.01	17.98	6 to 12.58	12			PTE: .10 NON: 1.05	PTE: - .06 NON: .49
OWSC/ Evasion	.09	2.94	1/54	5.14	2.75					PTE: - .24 NON: .10	PTE: - .28 NON: .15

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
For Criterion Variable: Fair: Student Rejoices (For Self)											
ASD/ Introversion	.07	3.46	1/51	5.66	24.51					PTE: - .08 NON: - .85	PTE: - .08 NON: - .41
SRI/ Reality	.03	5.08	1/51	8.59	14.88			19.15 to 30	18	PTE: - .03 NON: 1.27	PTE: - .06 NON: .37
OWSC/ Depression	.08	3.21	1/54	5.48	2.66					PTE: .09 NON: - .29	PTE: .24 NON: - .25
For Criterion Variable: Fair: Student Admits											
ASD/ Behavior	.10	2.83	1/51	5.00	13.67					PTE: .00 NON: - .91	PTE: - .01 NON: - .46
ASD/ Efficiency	.04	4.58	1/51	8.14	30.42	8 to 10.88	0	36.91 to 40	9	PTE: -1.23 NON: .96	PTE: - .33 NON: .25
SRI /Work	.08	3.22	1/51	5.79	15.64					PTE: - .09 NON: .89	PTE: - .06 NON: .42

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
OWSC/ Response Length	.09	2.89	1/54	4.83	6.70					PTE: 2.81 NON: - .12	PTE: .37 NON: - .07
For Criterion Variable: Fair: Teacher Does Solitary Work											
ASD/ Introversion	.01	7.87	1/51	13.14	23.53	8 to 17.61	25	33.07 to 40	5	PTE: - .88 NON: .35	PTE: - .38 NON: .37
ASD/ Attractiveness	.01	7.18	1/51	11.03	24.39	8 to 16.65	1	28.59 to 40	29	PTE: 1.53 NON: - .25	PTE: .51 NON: - .16
SRI/ Authority	.08	3.05	1/51	5.53	17.60					PTE: .87 NON: - .33	PTE: .27 NON: - .22
SRI/ Work	.05	3.95	1/51	6.72	15.05			22.33 to 30	11	PTE: .93 NON: - .12	PTE: .39 NON: - .11
For Criterion Variable: Profile of Learning Priorities (Student Teacher): Competent Management											
ASD/ Attitude	.01	6.23	1/54	9.74	36.47	8 to 33.01	24			PTE: 0 NON: 2.83	PTE: .01 NON: .50

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
SRI/ Self	.09	2.87	1/54	4.51	24.10	5.14 to 19.05	29			PTE: .36 NON: 1.51	PTE: .15 NON: .47
SRI/ Children	.05	3.90	1/54	6.55	23.40	6 to 20.90	18			PTE: - .96 NON: 1.91	PTE: - .26 NON: .27
OWSC/ Populars	.01	7.16	1/57	10.52	20.61	0 to 17.26	39	30.76 to 62	0	PTE: - .40 NON: .98	PTE: - .20 NON: .45
OWSC/ Evasion	.04	4.10	1/57	6.41	-.26			2.30 to 62	27	PTE: .27 NON: - .12	PTE: .34 NON: - .15
For Criterion Variable: Profile of learning Priorities(Student Teacher): Flexibility											
ASD/ Attitude	.01	7.76	1/55	12.31	35.35	8 to 31.85	13			PTE: .85 NON: 2.36	PTE: - .24 NON: .49
ASD/ Introversion	.09	2.98	1/55	5.06	22.04					PTE: - .19 NON: .55	PTE: - .10 NON: .39

Appendix B (cont.)

Interaction	Prob.	F	df	%Var.	Point of Intersect	Lower Region of significance bounds	%subjects	Upper Region of significance bounds	%subjects	Beta	Correl.
For Criterion Variable: Profile of Learning Priorities(Student Teacher): Professionalism											
ASD/ Attitude	.04	4.37	1/55	6.73	36.59	8 to 32.03	14			PTE: .48 NON: 2.74	PTE: .12 NON: .55
ASD/ Introversion	.03	4.60	1/55	7.57	24.55	8 to 17.02	25			PTE: - .53 NON: .41	PTE: - .28 NON: .28
SRI/ Children	.05	3.85	1/55	6.18	23.49	6 to 21.12	30			PTE: .22 NON: 2.91	PTE: .02 NON: .48
For Criterion Variable: Profile of Learning Priorities(Student Teacher): Responsibility											
ASD/ Introversion	.03	4.66	1/55	7.73	22.40	8 to 11.11	6			PTE: - .35 NON: .57	PTE: - .21 NON: .35
SRI/ Others	.08	3.12	1/55	5.33	19.32					PTE: .54 NON: - .98	PTE: .18 NON: - .29
JWSC/ Evasion	.04	4.29	1/58	6.79	0.93			4.41 to 62	9	PTE: .32 NON: - .03	PTE: .39 NON: - .04

References

Allen, D. W., & Mackin, R. A. Toward '76: A revolution in teacher education. Phi Delta Kappan, 1970, 51(9), 485-488.

Berliner, D. C., & Cahen, L. S. Trait-treatment interaction and learning. In F. Karlinger (Ed.), Review of research in education. Itasca, Illinois: Peacock, 1973.

Borich, G. D. Interactions among group regressions: Testing homogeneity of group regressions and plotting regions of significance. Educational and Psychological Measurement, 1971, 31, 251-253.

Borich, G. D. Homogeneity of slopes test for multiple regression equations with reference to aptitude treatment interactions. Journal of Experimental Education, 1972, 40, 39-42.

Borich, G. D. A recommended procedure for identifying trait-treatment interactions with special reference to post hoc analyses. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1973.

Borich, G. D., & Drezek, S. F. Evaluating instructional transactions. In G. D. Borich (Ed.), Evaluating educational programs and products. Englewood Cliffs, New Jersey: Educational Technology Publications, 1974.

Borich, G. D., & Godbout, R. C. Extreme groups designs and the calculation of statistical power. Educational and Psychological Measurement, 1974, 34(3), in press.

Borich, G. D., & Wunderlich, K. W. Johnson-Neyman revisited: Determining interactions among group regressions and plotting regions of significance in the case of two groups, two predictors, and one criterion. Educational and Psychological Measurement, 1973, 33(1), 155-159.

Bown, O. H., & Veldman, D. J. Scoring procedure and college freshman norms for the Self-Report Inventory. Research and Methodology Monographs, 1967, No. 1. (Austin, Texas: Research and Development Center for Teacher Education, The University of Texas)

Bracht, G. H. Experimental factors related to aptitude-treatment interactions. Review of Educational Research, 1971, 40, 627-645.

Bruner, J. The skill of relevance or the relevance of skills. Saturday Review, 1970, 53(14), 66-68.

Bruner, J. Relevance of education. New York: W. W. Norton, 1971.

Buchanan, M. M. Preparing teachers to be persons. Phi Delta Kappan, 1971, 52(10), 614-617.

Coan, R. W. Dimensions of psychological theory. American Psychologist, 1968, 23, 715-722.

Combs, A. W., et al. Florida studies in the helping professions. University of Florida Monographs, Social Studies, 1969, No. 37. (Gainesville, Florida: University of Florida Press)

Cronbach, L. J., & Snow, R. E. Aptitudes and instructional methods. (Preliminary ed.) Stanford, California: School of Education, Stanford University, 1973.

Cronbach, L. J., & Snow, R. E. Aptitudes and instructional methods. (Preliminary ed.) Stanford, California: School of Education, Stanford University, 1974.

Dinkmeyer, D. The C-group: Focus on self as instrument. Phi Delta Kappan, 1971, 52(10), 617-619.

Duncan, O. D. Path analysis: Sociological examples. The American Journal of Sociology, 1966, 12, 1-16.

Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart & Winston, 1968.

Emmer, E. T. Classroom Observation Scales. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1971.

Emmer, E. T., & Peck, R. F. Dimensions of classroom behavior. Journal of Educational Psychology, 1973, 64(2), 223-240.

Fruchter, B. Introduction to factor analysis. New York: D. VanNostrand, 1954.

Fuller, F. F. Concerns of teachers: A developmental conceptualization. American Educational Research Journal, 1969, 6, 207-226. (a)

Fuller, F. F. FAIR system manual. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1969. (b)

Fuller, F. F. A conceptual framework for a personalized teacher education program. Theory into Practice, 1974, 13(2), 112-122.

Fuller, F. F., & Newlove, B. W. Manual for use with "Meet your cooperating teacher." Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1969.

Fuller, F. F., Peck, R. F., Bown, O. H., Menaker, S. L., White, M. M., & Veldman, D. J. Effects of personalized feedback during teacher preparation on teacher personality and teaching behavior. Final Report. Project No. 5-0811, Personality, Teacher Education and Teaching Behavior Research Project. Report Series No. 4, Research and Development Center for Teacher Education, The University of Texas at Austin, 1969.

Gage, N. L. (Ed.) Handbook of research on teaching. Chicago: Rand McNally, 1963.

Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Ginott, H. Teacher and child. New York: Macmillan, 1972.

Goodlad, J. I. The schools vs. education. Saturday Review, 1969, 52(16), 59-61.

Goodman, P. Growing up absurd. New York: Random House, 1960.

Haak, R. A., Kleiber, D. A., & Peck, R. F. Student evaluation of teacher instrument, II, Manual. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1972.

Hall, G. E. The instrument for the analysis of science teaching: A system for measuring teaching behavior. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1969.

Hall, G. E. A manual for users of the IASTv2: A system of interaction analyses. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1972.

Herndon, J. How to survive in your native land. New York: Simon and Schuster, 1971.

Holt, J. How children fail. New York: Pittman, 1964.

Johnson, P. O., & Neyman, J. Tests of certain linear hypotheses and their applications to some educational problems. Statistical Research Memoirs, 1936, 1, 57-63.

Kozol, J. Death at an early age. Boston: Houghton-Mifflin, 1967.

Mandler, G., & Kessen, W. The language of psychology. New York: John Wiley, 1959.

Margenau, H. The nature of physical reality. New York: McGraw-Hill, 1950.

McNemar, Q. Psychological statistics. New York: Wiley, 1966.

Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

O'Neil, H. F., Jr. Effects of stress on state anxiety and performance in computer-assisted learning. Journal of Educational Psychology, 1972, 63(5), 473-481.

Ornstein, A. C. Systematizing teacher behavior research. Phi Delta Kappan, 1971, 52(9), 551-555.

Peck, R. F. Mental health in teacher education project. National Institutes of Mental Health, Grant No. 5T2MH-6635. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1958.

Peck, R. F. Promoting self-disciplined learning: A researchable revolution. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Peck, R. F. Personalized Teacher Education: Basic program plan. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1972. (a)

Peck, R. F. Profile of Learning Priorities, Form C. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1972. (b)

Potthoff, R. F. On the Johnson-Neyman technique and some extensions thereof. Psychometrika, 1964, 29, 241-256.

Rosenshine, B. Teaching behaviors and student achievement. New York: Fernhill, 1971.

Rosenshine, B., & Furst, N. Research in teacher performance criteria. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Royce, J. R. Factors as theoretical constructs. American Psychologist, 1963, 18(8), 522-528.

Schweiker, R. F. Factor scores aren't sacred: Comments on "Abuses of factor scores." American Educational Research Journal, 1967, 4, 168-170.

Silberman, C. E. Crisis in the classroom: The remaking of American education. New York: Random House, 1970.

Spielberger, C. D., Gorsuch, R. L., & Lishene, R. E. Manual for State-Trait Anxiety Inventory. Palo Alto, California: Consulting Psychologists Press, 1970.

Toffler, A. Future shock. New York: Random House, 1971.

Veldman, D. J. Fortran programming for the behavioral sciences. New York: Holt, Rinehart & Winston, 1967.

Veldman, D. J. Adjective Self-Description. Research and Methodology Monographs, 1970, No. 11R. (Austin, Texas: Research and Development Center for Teacher Education, The University of Texas)

Veldman, D. J. Automated sentence completion screening. Research and Methodology Monographs, 1971, No. 12. (Austin, Texas: Research and Development Center for Teacher Education, The University of Texas)

Veldman, D. J. Scoring procedures and report generators for COMPASS instruments. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1973.

Veldman, D. J., & Bown, O. H. Personality and performance characteristics associated with cigarette smoking among college freshmen. Journal of Consulting and Clinical Psychology, 1969, 33, 109-119.

Veldman, D. J., Menaker, S. L., & Peck, R. F. Computer scoring of sentence completion data. Behavioral Science, 1969, 14, 501-507.

Veldman, D. J., & Parker, G. V. C. Adjective rating scales for self description. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1968.

Walker, H., & Lev, J. Statistical inference. New York: Holt, 1953.

Watkins, J. E. Preliminary draft of a report on the development of TCCL (Teacher Concerns Checklist). Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1973.

Wehling, L. J., & Charters, W. W., Jr. Dimensions of teacher beliefs about the teaching process. American Educational Research Journal, 1969, 6, 7-30.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Yee, A. H., & Gage, N. L. Techniques for estimating the source and direction of causal influence in panel data. Psychological Bulletin, 1968, 70, 115-126.